Evolution of antimicrobial resistance in *Mycobacterium tuberculosis* studied in the field and the laboratory

INAUGURALDISSERTATION
zur
Erlangung der Würde eines Doktors der Philosophie

vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von
Sebastian Mark Gygli
aus Utzenstorf BE und Kanada

Basel, 2019

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von Herrn Prof. Dr. Sébastien Gagneux und Herrn Prof. Dr. Alex Hall.

Basel, den 11. Dezember 2018

Prof. Dr. Martin Spiess
Dekan

Summary

Human tuberculosis (TB), caused by the bacterium *M. tuberculosis*, is the leading cause of death due to an infectious agent. In 2017, there were an estimate of 10 million new cases and 1.6 million deaths caused by TB. The evolution of drug resistant *M. tuberculosis* variants is threatening disease control efforts. Whereas the treatment of drug-susceptible TB isolates involves a treatment regimen of 4 drugs which have to be taken on a daily basis for 6 months, treating drug resistant TB involves the administration of more toxic and expensive drugs for up to two years. Generating timely and accurate drug susceptibility testing results is essential for the administration of effective treatment. Effective treatment will result in fewer secondary cases generated per patient, as well as lower the probability of treatment failure and amplification of resistance. Drug resistance is primarily conferred by well-defined chromosomally encoded mutations, which offers the possibility of deducing antibiograms by sequencing the drug-resistance related target genes and detecting mutations therein. In Chapter 4, we assessed the utility of whole genome sequencing in predicting accurate drug susceptibility profiles, as well as predicting quantitative levels of drug resistance. For this, we compared whole genome sequences with results of two different methods suitable in determining quantitative levels of drug resistance. We concluded that whole genome sequencing is highly effective in predicting drug resistance profiles and is in part able to predict quantitative levels of drug resistance.

It has generally been assumed that drug resistant *M. tuberculosis* variants do not transmit efficiently, as drug resistance comes at a cost. Drug resistance-conferring mutations often target essential, highly conserved genes, causing a physiological cost, resulting in a reduced number of secondary cases generated. However, these costs may be mitigated by secondary, so-called compensatory mutations. Examples of compensatory evolution in *M. tuberculosis* are limited to a few examples. In chapter 6 we attempted to identify biological pathways mutated more frequently than expected by chance in multidrug-resistant (MDR: resistance against rifampicin (RIF) and isoniazid (INH)) *M. tuberculosis* strains which generated secondary cases, compared to strains which did not transmit frequently. We identified nucleotide biosynthesis pathways as more frequently mutated than expected

by chance in strains that transmit frequently, potentially hinting at a connection to RIF-resistance and altered nucleotide requirements of RIF-resistant strains. Although only 560000 cases of rifampicin resistant *M. tuberculosis* infections were reported in 2017, the incidence is not equally distributed globally. For instance, countries of the former Soviet Union are heavily affected by drug resistant variants for reasons not well understood. In Chapter 5, we analysed a population-based, nation-wide collection of clinical MDR *M. tuberculosis* strains, isolated between 2011 and 2013 in Georgia. We identified Georgian prisons as a hotspot for transmission of MDR TB. Furthermore, we were able to demonstrate that MDR strains belonging to lineage 2 transmit more frequently compared to lineage 4 strains. Strains harbouring compensatory mutations transmitted more frequently than non-compensated strains. Furthermore, strains connected directly or indirectly to prisons were more likely to harbour compensatory mutations, indicating that prisons act as incubator for highly transmissible *M. tuberculosis* strains. In total, 41 % of all transmission events of MDR-TB isolates were directly or indirectly associated with incarceration.

Since the discovery that single drug regimens for the treatment of *M. tuberculosis* quickly lead to resistance, TB is treated with combination therapies. The current WHO-endorsed regimen for drug susceptible TB contains four drugs, including RIF and INH. Patient non-adherence and pharmakogenomic variation among patients have been demonstrated to generate sub inhibitory drug concentrations, which in turn are implicated in the evolution of drug resistance. In chapter 7 we investigate the influence of sub inhibitory concentrations of RIF and INH alone and in combination. Out data hints at the possibility that katG mutations, conferring INH resistance, are also involved in conferring low-level RIF resistance. This observation might explain the observation why, compared to INH mono-resistance, RIF mono-resistance is rare in clinico and can be used as a surrogate marker for multidrug-resistant TB.

Contents

Acknowledgements

The completion of this thesis would not have been possible without the help and support of many people. I am deeply grateful to have been able to work together with so many wonderful people at the Swiss Tropical and Public Health Institute and the National Center for Tuberculosis and Lung Diseases in Tbilisi, Georgia.

I am indebted to my Doktorvater Prof. Dr. Sébastien Gagneux, who gave me a lot of freedom to pursue my interests in evolutionary biology/infection biology, even if that meant performing slightly unusual experiments with dubious chances for success. Sébastien, I am thankful for all the support you gave me during the past years and for your dedication, enthusiasm, guidance and mentorship. You inspired me and exerted a great influence on me as a fledging scientist. Besides your personal mentoring, you created a wonderful work environment which greatly contributed to me enjoying my stay as a member of your group.

I want to thank Prof. Dr. Alex Hall for being on my thesis committee and providing help and guidance from the very beginning of my doctoral studies. I would also like to thank Prof. Dr. Pascal Mäser for agreeing to chair my thesis defense.

I want to extend my gratitude to Levan Jugheli, Rusudan Aspindzelashvili and Nino Maghradze for performing the incredible work of organizing and generating the raw material from literally thousands of TB isolates, which served as the basis for two chapters of my thesis. Without your hard work, a great deal of my thesis would not have been possible. I would also like to thank Levan in particular for hosting me during my visit to Georgia in 2016, which I thoroughly enjoyed. None of the work presented in this thesis would have been possible without the support of my work colleagues and friends. I am especially grateful to my friends: I want to thank Chloé Loiseau for helping me out uncountable times and for providing me with the means to efficiently perform many analyses which are at the core of my thesis. Without your continued support I would have gotten nowhere. I want to thank Monica Ticlla for the delightful collaboration on chapter 6, which produced truly exciting results. I am very grateful for your patience, technical ex-

pertise and support. I want to thank Rhastin Castro for the great time we had pipetting clumps of bacteria in 36 h intervals. I want to thank Liliana Ruthaiwa for all the fun discussion we had over some Schneider in Leiden, Wädenswil and Basel. Thank you all for the laughter, fun evenings and outings I was able to enjoy with you.

I want to thank Andrej Trauner for all his support from the very beginning of my doctoral studies and for being a role model for my scientific endeavours. I greatly appreciated your wit, critical thinking and help. Lab work was a substantial part of my thesis and it would not have worked without the help of Sonia Borrell who trained me to work in the BSL3 and who continuously supported all my, at times, odd experimental undertakings. I would also like to thank Julia Feldmann and Miriam Reinhard for their incredible support, especially concerning genome sequencing and the experimental evolution experiment. I would like to thank David Stucki, Mireia Coscolla and Daniela Brites for their support when I was starting out with my genomics analysis and scripting. I am grateful to all the other members of the TB research group *sensu stricto et lato*: Ainhoa Arbues Arribas, Peter Major, Jerry Hella, Fabrizio Menardo and Isaak Otchere for their friendship and support. I also want to extend my gratitude to Severin Gsponer and Phillipp Schmid for the stimulating discussions on so many different topics.

I am deeply grateful for the patience, love and support from Alexandra Mushegian. Sasha, I am thankful to have you in my life. Thank you for being here for me during the good and the few challenging times (the dead crab she said, she said!). I would like to express my sincerest gratitude to my parents Albin and Yvonne Gygli and to my brother Albin Derek Gygli, for their continued support throughout my whole life. Thank you for making me the person I am and for getting me where I am today. Nothing would have been possible without you.

1. Introduction

Human tuberculosis (TB), caused by the bacterium *Mycobacterium tuberculosis*, is by far the deadliest infectious disease in the history of the human species, having claimed more lives than any other pathogen, bacterial or viral. In the past 200 years, TB has claimed the lives of at least 1 billion people (Paulson, 2013). In 2017, there were an estimated 10 million new cases of TB and an estimated 1.6 million deaths caused by TB (World Health Organization 2018a). A number of factors have contributed to the success of the pathogen. First, *M. tuberculosis* is transmitted aerogenously via bacteria-containing droplet nuclei and is believed to have a low infectious dose of 1-200 cells (Sakamoto, 2012), greatly facilitating host-to-host transmission. Furthermore, TB infections lead to chronic disease which, in absence of effective treatment, can last for years, during which the host may spread the disease (Behr *et al.*, 2018). Cleared TB infections do not lead to the development of long-lasting protective immunity, resulting in possible multiple infections during the lifetime of a person in high-incidence settings (Shen *et al.*, 2006). However, global incidence of TB has been dropping continuously since accurate records on incidence have been collated (Wilson, 1990). The discovery of the tubercle bacillus by Robert Koch in 1882 (Koch, 1882) paved the way for implementing effective public health interventions, especially isolation of TB patients (Zürcher *et al.*, 2016). In the Western hemisphere, TB incidence has dramatically declined, to the extent of rendering TB an anecdotal disease; which is linked to the reduction of transmission possibilities. Improved living conditions lead to less crowding, which reduces the exposure of, for example, family members living together with a person infected with TB. The discovery of antimicrobial compounds active against *M. tuberculosis* helped to speed up the decline in TB incidence in the Western world. Early setbacks due to the evolution of drug resistance when using a single drug to treat TB were soon circumvented by the establishment of combination therapy (Dickinson *et al.*, 1977). However, the TB epidemic is far from contained. Especially the emergence and spread of the human immunodeficiency virus (HIV) in population-rich regions causes frequent and detrimental HIV-TB co-infections (Brites *et al.*, 2012) and poses a sub-

stantial challenge to reducing TB mortality and morbidity (World Health Organization, 2018a).

1.1. Treatment of TB and the evolution of drug resistance

While TB is a curable disease, doing so it is not trivial. At least 4 active drugs are required to treat drug-susceptible TB, and these drugs need to be taken for 6 months on a daily basis without interruption. The long treatment duration, as well as the high cost of treating drug-resistant TB (Marks et al. 2014) poses considerable challenges for weak health systems (Atun *et al.*, 2010). The current treatment standard, directly observed therapy short course (DOTS), widely implemented in the last decade of the 20$^{\text{th}}$ century, aims to ensure that patients take the prescribed drugs on a regular basis without missing doses (World Health Organization, 1997). However, patients can fail to adhere to treatments for a multitude of reasons, including severe side effects of antituberculous drugs, lack of financial capabilities to access drug treatment, and drug unavailability. Interruption of treatment is therefore to be expected in a certain proportion of patients (Munro *et al.*, 2007). Interrupted treatment is a risk factor for developing drug resistance. However, drug resistant strains have been demonstrated even in patients who take the prescribed drugs on a regular basis (Calver *et al.*, 2010). There are several factors thought to contribute to the evolution of drug-resistant strains within patients. Individual variation in drug metabolism can result in fixed-dose regimens not providing sterilizing concentrations in each patient for the required period of time (Srivastava *et al.*, 2011). Pharmacogenomic differences, as well as treatment interruption lead to sub-inhibitory drug concentrations, which are known to facilitate the evolution of drug-resistant variants (Andersson *et al.*, 2014). In Chapter 7, we investigate how drug environments below the minimal inhibitory concentration (sub-MIC) may select for MDR phenotypes. There is preliminary evidence that combinations of sub-MIC RIF and INH concentrations effectively select for MDR TB strains.

Limiting the evolution of drug resistance requires the acquisition of timely and accurate drug resistance profiles. Undetected resistance may lead to the administration of less than four effective antituberculous drugs, and in many cases the consequence will be treatment failure and the amplification of resistance (Zhang *et al.*, 2015a). The current gold standard for drug susceptibility testing in TB involves the assessment of bacterial growth at defined concentrations, testing growth of patient isolates at defined drug con-

centrations to determine if a strain is deemed clinically resistant. The reliance on these time-consuming tests may have slowed the investigation of distributions of minimal inhibitory concentrations of resistant strains. There is considerable variability between the levels of resistance conferred by drug resistance mutations. Some resistance mutations do not elevate the MIC far beyond the critical concentration; strains carrying these mutations may still be treated with increased doses of the drug. In light of the limited set of drugs active against *M. tuberculosis*, quantifying the distribution of MICs conferred by different drug resistance mutations warrants more investigation. As reviewed in Chapter 3, drug resistance in M, tuberculosis is thought to be conferred predominantly by chromosomal mutations. This makes detection of resistance mutations using whole genome sequencing (WGS) a powerful tool in predicting drug susceptibility testing (DST) profiles. However, the utility of WGS to accurately predict DST profiles relies on being able to correlate identified mutations with high-quality drug susceptibility testing results, which are still rare. In Chapter 4, we examined how well WGS is able to predict DST profiles and quantitative levels of drug resistance after filtering for known phylogenetic markers. We concluded that WGS has great potential to replace culture-based DST in the future, and discovered that there is potential room for improvement in treatment regimens, as not all drug resistance mutations render strains resistant to drug concentrations beyond the therapeutic window of a drug.

1.2. Control of the TB epidemic and prisons

It is unlikely that the TB epidemic is going to be ended with antibiotics alone. Even in highly developed health systems, there are still cases among the resident population (World Health Organization, 2018a). TB is able to enter a latent phase, which generates a currently undetectable reservoir from which new cases can continuously emerge and generate secondary cases, which in turn may enter dormancy again, restarting the cycle of hide and seek (Jassal *et al.*, 2010). Ending the TB epidemic will require a highly effective vaccine that generates long-lasting protective immunity, which is currently not in sight (Kaufmann *et al.*, 2014). Until an effective vaccine is developed, public health interventions, including active case-finding campaigns, will remain important to reduce TB incidence. Although the benefits of active case-finding campaigns on the population at large remain unclear (Kranzer *et al.*, 2013), active screening is able to reduce the burden of TB in "stationary" high risk populations, e.g. prisoners (World Health Organization *et al.*, 2018).

TB incidence rates among prisoners are generally higher than in the general public (Baussano *et al.*, 2010) due to a number of factors facilitating the spread of the disease. For instance, in many countries prisoners are held in confined, overcrowded spaces with poor ventilation. Prolonged exposure to an active case greatly increases the probability of contracting the disease (Reichler *et al.*, 2018). Furthermore, prisoners frequently experience difficulties in accessing adequate medical services (Department of State of the United States of America, 2012). Active case-finding campaigns in prisons should be highly effective in quickly identifying active cases and enabling the timely administration of effective treatment. However, despite the relative ease of identifying TB cases in prisons, the discrepancy in incidence rates between prisoners compared to the general public is shocking. For instance, in 2013, the incidence of TB among prisoners in Georgia prisoners was 25.5 times higher than that in the general population, despite active case finding campaigns being in place (World Health Organization, 2014b). Prisoners are also at a dramatically increased risk of contracting drug-resistant strains (Aerts *et al.*, 2000; Stuckler *et al.*, 2008; Kenyon, 2009), again demonstrating the partial failure of TB control programs. Furthermore, there is a correlation between the proportion of incarcerated people and the incidence of drug-resistant TB in the general population (Stuckler *et al.*, 2008). The impact of incarceration on the TB epidemic in the general public in countries of the former Soviet Union has not received sufficient attention. For instance, in the case of the Republic of Georgia, there are no programs in place to ensure that prisoners continue to be treated for TB after their release from prison. In 2012, following the exposure of appalling human rights violations in Georgian prisons, nearly 10000 prisoners were released from prison. In the absence of effective follow-up, the mass amnesty is very likely to have had a profound impact on the epidemic of drug-susceptible and multidrug-resistant TB. Investigating the reasons of why prisoners in Georgia appear to be at a great disadvantage in terms of being TB-free is beyond the scope of this thesis. However, in Chapter 5, we were able to show that neglecting the prison population has direct consequences for the general public, as apparent from spillover events of highly transmissible TB strains from prisons into the general public. We hope that such findings will invigorate programs to reduce TB incidence in prisons and limit the spillover to the general public.

1.3. Bacterial factors and transmissibility of drug resistant TB

While improvements in DST and public health programs will reduce TB burden, designing effective treatment regiments requires a thorough understanding of resistance and compensation mechanisms. The great majority of antibiotics directly target essential metabolic processes in bacteria (Chapter 3). Acquiring resistance to the drugs requires mutating the genes encoding the drug target. However, these essential proteins are highly conserved, and non-synonymous mutations frequently cause disruptions in the bacteria's physiology. The magnitude of drug resistance- induced disruption of physiological processes, manifesting as a reduced *in vitro* growth rate, is known to depend on the resistance mutation and the genetic background in *M. tuberculosis* and manifests itself as a reduced *in vitro* growth rate (Gagneux *et al.*, 2006b). The reduced growth may cause the drug-resistant strain to generate fewer secondary cases compared to drug- susceptible strains in absence of the drug. Secondary, so-called compensatory mutations, not directly involved in conferring drug resistance, are able to mitigate these drug-resistance related fitness costs and restore the capability of drug resistant strains to efficiently generate secondary cases (Chapter 5, (Merker *et al.*, 2018)). Furthermore, the concomitant presence of multiple drug resistance mutations affecting essential genes seems to generate a multitude of targets of compensatory evolution (Moura de Sousa *et al.*, 2017). We do not have a good understanding of the molecular mechanisms underlying the measurable fitness deficit and its compensation observed in drug resistant *M. tuberculosis* strains. The study of population-based collections may help in elucidating what the targets of compensatory evolution are, by analysing what mutations strains transmitting frequently acquire. Combined with metabolomic, proteomic and transcriptomic data, this will greatly enhance our understanding of fitness cost compensation in *M. tuberculosis* by linking genotypic with phenotypic readouts. In Chapter 6 we present work in trying to identify genetic factors involved in fitness cost compensation by analysing a population-based MDR-TB strain collection by detecting which mutations appear in strains that transmit frequently. These data will serve as a first step in elucidating how *M. tuberculosis* strains are capable of acquiring resistance to an numerous drugs via chromosomal mutations in essential genes, yet retaining the capability to generate secondary cases (Chapter 5, (Shah *et al.*, 2017)).

2. Objectives and outline

2.1. Thesis aims

This thesis aims at addressing multiple aspects concerning drug resistance in *M. tuberculosis*, ranging from inferring drug resistance profiles from whole genome sequences, the evolution of drug resistance under sub inhibitory drug concentrations, the transmissibility of multidrug resistant strains and the detection of mechanisms rendering drug resistant variants transmissible.

2.1.1. Specific objectives

- **Objective 1** Review the current knowledge on drug resistance mechanisms and the evolution of acquired drug resistance in *M. tuberculosis* (Chapter 3).

- **Objective 2** Assess the utility of whole genome sequencing to infer accurate drug resistance profiles in clinical *M. tuberculosis* isolates (Chapter 4).

- **Objective 3** Study transmission of multidrug-resistant *M. tuberculosis* strains in the Republic of Georgia using a population-based collection *M. tuberculosis* strain and infer bacterial, host and environmental factors associated with successful transmission (Chapter 5).

- **Objective 4** Identify mutations or pathways under selection in transmissible multidrug-resistant *M. tuberculosis* strains using the population-based collection of *M. tuberculosis* strains from Georgia (Chapter 6).

- **Objective 5** Study the evolution of mono- and multidrug-resistance in clinical *M. tuberculosis* strains under selection of sub-inhibitory drug concentrations (Chapter 7).

2.2. Outline

In *Chapter 3*, we reviewed mechanisms of intrinsic and acquired drug resistance in *M. tuberculosis* and attempted to shed light on acquired drug resistance from an evolutionary perspective.

In *Chapter 4* we assess the utility of whole genome sequencing for the prediction of quantitative levels of drug resistance in a diverse set of 176 clinical *M. tuberculosis* isolates by comparing the results of two phenotypic drug susceptibility testing methods with the genotype inferred from the whole genome sequences.

Chapter 5 aims at studying the transmission of highly drug-resistant *M. tuberculosis* isolates and gaining insight into the epidemic of multidrug-resistant *M. tuberculosis* in the former Soviet Union. For this we a subjected a population-based collection of multidrug-resistant *M. tuberculosis* strains from Georgia collected between 2011 and 2013 to whole genome sequencing to infer transmission networks. Based on the transmission networks, we analysed bacterial, environmental and host factors for their association with successful transmission. *Chapter 6* is concerned with identifying mechanisms involved in mitigating drug resistance-related fitness costs, enabling the drug resistant *M. tuberculosis* strains to efficiently generate secondary cases. For this, we analysed the transmission networks inferred in Chapter 5, identified mutations private to strains in transmission clusters and subjected the mutations to network analysis-based pathway enrichment.

Chapter 7 investigates the effect of sub inhibitory drug concentrations of rifampicin and isoniazid alone or in combination on the evolution of drug resistance. For this we subjected three clinical *M. tuberculosis* strains with differing genetic backgrounds to serial passage for approximately 180 bacterial generations. Drug concentrations increased from 1/64 to 4 x of the minimal inhibitory concentration. We subjected a subset of the surviving endpoint cultures to whole genome sequencing to detect mutations associated with drug resistance.

In *Chapter 8*, we summarise the findings of this thesis and puts them into context of the current state of knowledge, while addressing remaining challenges and identifying potential directions of future research.

3. Evolution of drug resistance in *Mycobacterium tuberculosis*: Mechanistic and evolutionary perspectives

Sebastian M. Gygli[1,2], Sonia Borrell[1,2], Andrej Trauner[1,2], Sebastien Gagneux[1,2]

[1]Swiss Tropical and Public Health Institute, Basel, Switzerland
[2]University of Basel, Basel, Switzerland

This chapter has been published in FEMS Microbiology Reviews, Volume 41, Issue 3, 1 May 2017, Pages 354–373

3.1. Abstract

Antibiotic-resistant *Mycobacterium tuberculosis* strains are threatening progress in containing the global tuberculosis epidemic. *M. tuberculosis* is intrinsically resistant to many antibiotics, limiting the number of compounds available for treatment. This intrinsic resistance is due to a number of mechanisms including a thick, waxy, hydrophobic cell envelope and the presence of drug degrading and modifying enzymes. Resistance to the drugs which are active against *M. tuberculosis* is, in the absence of horizontally transferred resistance determinants, conferred by chromosomal mutations. These chromosomal mutations may confer drug resistance via modification or overexpression of the drug target, as well as by prevention of prodrug activation. Drug resistance mutations may have pleiotropic effects leading to a reduction in the bacterium's fitness, quantifiable e.g. by a reduction in the *in vitro* growth rate. Secondary so-called compensatory mutations, not involved in conferring resistance, can ameliorate the fitness cost by interacting epistatically with the resistance mutation. Although the genetic diversity of *M. tuberculosis* is low, compared to other pathogenic bacteria, the strain genetic background has been demonstrated to influence multiple aspects in the evolution of drug resistance. The rate of resistance evolution and the fitness costs of drug resistance mutations may vary as a function of the genetic background.

3.2. Introduction

Human tuberculosis (TB), a devastating disease caused by the gram-positive, acid-fast eubacterium *Mycobacterium tuberculosis*, was classified as a global health emergency by the World Health Organization in 1993. TB remains one of the deadliest infectious diseases with an estimated 1.8 million deaths occurring per year, mainly in the developing world (World Health Organization, 2016a). Although the incidence of TB has declined drastically over the past decades, there were an estimated 10.4 million new cases in 2015, of which 0.48 million were caused by *M. tuberculosis* strains classified as multidrug-resistant (MDR – resistant to the first-line drugs rifampicin and isoniazid). Drug-resistant *M. tuberculosis* strains are a major global health concern because treatment of these cases requires second-line drugs, which are less effective, more expensive and more toxic, as well as sophisticated infrastructure for drug susceptibility testing not readily available in resource-limited settings. TB treatment success rates of cases caused by MDR/XDR variants of *M. tuberculosis* are alarmingly low, with only 54 % of MDR and 28 % of extensively drug-resistant (XDR – MDR plus resistance to fluoroquinolones and any second-line injectable aminoglycoside / cyclic peptide) cases resulting in cure, compared to 83 % of drug-susceptible cases (World Health Organization, 2016a). For clarity we only refer to MDR/XDR *M. tuberculosis* variants when the defined resistance profiles are meant. Otherwise we use the term "drug resistance" to refer to the topic in general, irrespective of specific drug resistance profiles.

The first effective antituberculous drug, streptomycin, was discovered in 1944 (Schatz *et al.*, 1944). The newly discovered drug was immediately used for treatment of TB patients. The condition of many individual TB patients receiving streptomycin improved during the first months of treatment, only to then deteriorate again as treatment continued. It was soon understood that this was due to the evolution of resistant *M. tuberculosis* strains, rendering streptomycin ineffective (Crofton *et al.*, 1948). To limit the evolution of resistance, the British Medical Research Council pioneered the first combination therapy for the treatment of a disease by using para-aminosalicylic acid (Lehmann, 1946) together with streptomycin for treatment of pulmonary TB (Medical Research Council, 1950). The subsequent years saw the introduction of an array of different antituberculous drugs. The discovery of rifampicin in 1965 (Sensi, 1983) and the subsequent use of the drug in TB treatment was a game-changer, allowing dramatically shortened treatment duration from 18 months or more to 9 months (British Thoracic and Tuberculosis Association, 1975). During the 1990s, the current standard 6-month regimen known as Directly Observed Therapy Short Course (DOTS) was introduced by the World Health Organization (World

Health Organization, 1997). This regimen consists of 2 months treatment with isoniazid, rifampicin, ethambutol and pyrazinamide followed by 4 months of isoniazid and rifampicin (World Health Organization, 1991), and is highly effective for drug-susceptible TB (Feng-Zeng *et al.*, 1996; Frieden *et al.*, 1995).

A short treatment duration and reduction of adverse drug effects are crucial for increasing patient treatment adherence, which is known to influence the evolution of drug resistance (Mahmoudi *et al.*, 1993). However, despite the early establishment of TB combination therapies showing high cure and low relapse rates (British Thoracic and Tuberculosis Association, 1975), drug-resistant *M. tuberculosis* strains continued to evolve in both high and low incidence settings. MDR *M. tuberculosis* variants evolved on multiple occasions in different parts of the world (Eldholm *et al.*, 2015; Cohen *et al.*, 2015). Furthermore, differences in the quality of public health systems contributed to the spread of drug-resistant *M. tuberculosis* variants leading to the unequal distribution of incidence rates of drug-resistant variants around the world we observe today (World Health Organization, 2016a). In the absence of an effective vaccine (Kaufmann *et al.*, 2014), there is an urgent need for new treatment regimens, drugs and diagnostics to slow the evolution of drug resistance and limit transmission of resistant variants, as well as to ameliorate the treatment outcome of patients infected with MDR/XDR *M. tuberculosis* strains. Understanding the molecular mechanisms and the evolutionary trajectory of drug resistance is important to limit the *de novo* evolution and subsequent spread of resistant *M. tuberculosis* strains. The first part of this review will summarise intrinsic and acquired mechanisms of drug resistance in *M. tuberculosis*; these are analysed in more detail in several recently published reviews (Nguyen, 2016; Zhang *et al.*, 2015b; Smith *et al.*, 2013; Nash, 2016). The second part of this review will focus on our current understanding of the evolutionary biology of drug resistance in *M. tuberculosis*.

3.3. Mechanisms of drug resistance in *Mycobacterium tuberculosis*

Members of the genus *Mycobacterium* have long been noted for their intrinsic resistance to a wide array of antibiotics. This has mainly been attributed to the unusually thick, lipid-rich cell envelope (Jarlier *et al.*, 1994). After penetrating the cell envelope, certain antibiotics may be cleaved enzymatically or altered structurally to render them ineffective (Wang *et al.*, 2006; Chambers *et al.*, 1995; Warrier *et al.*, 2016). Furthermore, a number of efflux systems have been identified in *M. tuberculosis*, but their significance in confer-

ring clinically relevant levels of drug resistance is a matter of debate. As efflux systems have been observed to be expressed under varying conditions (Li *et al.*, 2015; Adams *et al.*, 2011; Gupta *et al.*, 2010), they might serve as a stepping stone for high-level drug resistance. A further peculiarity of *M. tuberculosis* is the apparent absence of ongoing horizontal gene transfer (Cole *et al.*, 1998; Bolotin *et al.*, 2015; Gagneux *et al.*, 2007). Although there have been reports of horizontal gene transfer between "species" of the genus *Mycobacterium* (Rabello *et al.*, 2012), horizontal gene transfer does not seem to be a driving factor in the acquisition of antimicrobial resistance in *M. tuberculosis*. The vast majority of drug resistance phenotypes in *M. tuberculosis* can be explained by chromosomal mutations and not by resistance plasmids or other mobile genetic elements. To study the mechanisms of drug resistance, many studies have been performed on a multitude of different mycobacterial species, due to the often lower pathogenicity/biosafety requirements and faster growth properties of these mycobacteria compared to *M. tuberculosis*. The most widely used model is *Mycobacterium smegmatis*, an environmental mycobacterium with a genome roughly 1.5x the size of that of *M. tuberculosis*. **We should therefore be cautious in applying the results of these studies directly to *M. tuberculosis*.**

3.4. Intrinsic drug resistance in *M. tuberculosis*

3.4.1. The mycobacterial cell wall & drug penetration

The intrinsic resistance of mycobacteria against several classes of antibiotics has commonly been attributed to the unusual composition and structure of the mycobacterial cell envelope. Compared to other gram-positive bacteria, the cell wall of members of the genus *Mycobacterium* is much thicker and more hydrophobic, due to the presence of a wide array of different lipids that include mycolic acids. Many studies (reviewed in (Jarlier *et al.*, 1994; Brennan *et al.*, 1995; Nguyen *et al.*, 2009; Sarathy *et al.*, 2012) performed in different mycobacterial species demonstrated that the composition of the cell envelope and the low numbers of porins (Mailaender *et al.*, 2004) contribute significantly to the cell envelope's low compound permeability. A major constituent of the cell wall is a layer of lipids, which are covalently linked to the peptidoglycan layer via arabinogalactan. Furthermore, the cell wall contains "extractable" immunogenic glycolipids (Brennan *et al.*, 1995). The lipid-rich nature renders the cell wall extremely hydrophobic and prevents the permeation of hydrophilic compounds. It is thought that small hydrophilic compounds, including many antibiotics active against *M. tuberculosis*, can only traverse

the cell wall via water-filled porins. Heterologous expression of the *M. smegmatis* porin MspA in *M. tuberculosis* did indeed decrease the minimal inhibitory concentration for several hydrophilic drugs (Mailaender *et al.*, 2004), indicating that porins might play an important role in the diffusion of hydrophilic antibiotics across the cell wall of *M. tuberculosis*. However, until recently reports on the presence of porins in *M. tuberculosis* were lacking. The outer membrane channel protein CpnT was demonstrated to be involved in nutrient uptake in *Mycobacterium bovis* BCG and *M. tuberculosis* (Danilchanka *et al.*, 2014) and in mediating susceptibility to nitric oxide and antibiotics in *Mycobacterium bovis* BCG (Danilchanka *et al.*, 2015). CpnT seems to be under positive selection in clinical *M. tuberculosis* isolates, demonstrated by the overrepresentation of non-synonymous mutations in the gene encoding CpnT (Rv3903c). However, the role of CpnT in mediating drug susceptibility to hydrophilic antibiotic compounds in *M. tuberculosis* needs further investigation, as CpnT deletion mutants do not demonstrate drug resistance phenotypes *in vitro*. However, the studies confirm the presence of porins in the outer membrane of *M. tuberculosis* and their role in uptake of small hydrophilic compounds. Furthermore, the physical organization of the cell wall lipids is believed to limit the membrane's fluidity. A recent study (Rodriguez-Rivera *et al.*, 2017) assessed the membrane fluidity in live cells of *M. smegmatis* and other actinobacteria by measuring the reorganisation of fluorescein-labelled therealose analogs by mycolyltransferases. The study demonstrated that, compared to other actinobacteria, *M. smegmatis* has the lowest membrane fluidity. This is thought to be a function of mycolic acid structure (length & presence of functional groups). Interestingly, exposure of *M. smegmatis* to sub-inhibitory concentrations of ethambutol increases the membrane's fluidity and diffusion of compounds across the cell envelope. This offers the possibility for novel drug combination therapies, as the reduction of the membrane fluidity using ethambutol can render *M. tuberculosis* susceptible against drug classes it is normally resistant against (Abate *et al.*, 1997; Bosne-David, 2000).

The peculiar characteristics of the mycobacterial cell envelope hinder the diffusion of hydrophobic molecules including members of several antibiotics belonging to the classes of macrolides, rifamycins, tetracyclines and fluoroquinolones (Brennan *et al.*, 1995). However, it does appear that the rate of diffusion is a function of molecule hydrophobicity to a certain extent, with hydrophobic molecules diffusing more readily through the mycobacterial cell envelope (Liu *et al.*, 1999; Rastogi *et al.*, 1990). The hypothesis that the cell envelope lipids are a major factor in the intrinsic resistance of mycobacteria to many hydrophobic antibiotics is further substantiated by studies performed with mutants defective in lipid synthesis, which are susceptible to drugs that the corresponding wild

type strain is resistant against (Liu *et al.*, 1999). A recent study modelled the permeation of compounds through the mycobacterial cell wall and demonstrated that lipophilicity is an important but not exclusive factor of compound permeability (Janardhan *et al.*, 2016).

3.4.2. Drug inactivation by *M. tuberculosis*

After penetrating the cell wall as an initial defence layer, antibiotics may be cleaved enzymatically to render them ineffective. One of the most prominent examples is the enzymatic degradation of β-lactam antibiotics by β-lactamases, which hydrolyse the β-lactam ring of the antibiotics. Early studies involving penicillin demonstrated that *M. tuberculosis* is intrinsically resistant to this class of antibiotics (Abraham *et al.*, 1941). The genome of *M. tuberculosis* encodes a single class A β-lactamase termed BlaC thought to localize to the periplasmatic space, either anchored in the outer leaflet of the plasma membrane as a lipoprotein or unbound. The *M. tuberculosis* β-lactamase shows broad substrate specificity (including carbapenems), albeit with varying affinities, and is considered an extended-spectrum β-lactamase. BlaC is irreversibly inhibited by the β-lactamase inhibitor clavulanate (Wang *et al.*, 2006; Hugonnet *et al.*, 2007). Due to the increasing numbers of cases caused by MDR/XDR *M. tuberculosis* strains, there has been a renewed interest in the use of β-lactam antibiotics in the treatment of TB. An early, small study reported no beneficial effect of including an amoxicillin/clavulanate combination in a salvage regimen (a regimen of last resort with unproven efficacy) to treat patients infected with MDR *M. tuberculosis* strains (Yew *et al.*, 1995). Since then, several *in vitro* (Chambers *et al.*, 1995; Hugonnet *et al.*, 2009) and *in vivo* studies (Payen *et al.*, 2012; De Lorenzo *et al.*, 2013) reported encouraging results on treatment outcomes with various regimens by including β-lactam antibiotics with clavulanate. However, some MDR/XDR *M. tuberculosis* isolates still appear to be resistant to mereponem/clavulanate or amoxcicillin/clavulanate without harboring any mutations that could explain the observed variability in susceptibility to these drugs (Cohen *et al.*, 2016). The true value of β-lactam antibiotics for the treatment of drug-resistant *M. tuberculosis* variants still needs further assessment. Given the positive results in diverse studies, the demonstrated safety profile of β-lactam antibiotics/β-lactamase inhibitors and the limited treatment options for MDR/XDR TB, warrants further investigation into treatment regimens including this class of antibiotics.

Apart from drug cleavage, antibiotics may be inactivated by modification, e.g. by methylation or acetylation. To date, the best described mechanism of drug inactiva-

tion by chemical modification in *M. tuberculosis* is the acetylation of various aminoglycoside/cyclic peptide antibiotics used for the treatment of MDR TB by the enhanced intracellular survival protein (Eis). Eis has been demonstrated to acetylate and inactivate the clinically relevant second-line injectable aminoglycoside antibiotic kanamycin A (Zaunbrecher *et al.*, 2009), as well as the cyclic peptide antibiotic capreomycin (Houghton *et al.*, 2013). Several promotor mutations identified in clinical *M. tuberculosis* isolates lead to overexpression of Eis, which in turn confers low-level resistance against kanamycin A but not amikacin (Zaunbrecher *et al.*, 2009; Kambli *et al.*, 2016). It is not clear if Eis overexpression alone leads to clinically relevant levels of capreomycin resistance (Kambli *et al.*, 2016). Overexpression of Eis therefore might serve as a stepping stone for the evolution of high-level aminoglycoside/cyclic peptide resistance. Recently, a novel mechanism of drug inactivation was discovered in *M. tuberculosis*. The pyrido-benzimidazole compound "14" was described as having potent bactericidal activity against aerobically growing *M. tuberculosis* (Warrier *et al.*, 2015). Compound 14 may be N-methylated by a previously unknown methyltransferase encoded by the gene Rv0560c. The methylated compound 14 is unable to inhibit its target, the decaprenylphosphoryl-β-D-ribose 2-oxidase (DprE1), which is involved in arabinogalactan synthesis (Warrier *et al.*, 2016). Although this is a novel mechanism of drug resistance in *M. tuberculosis*, and in bacteria in general, it has no known clinical relevance to date.

3.4.3. Enzymatic drug target modification

Many antibiotics in use are natural products produced by bacteria, which requires the producing bacteria to be resistant to these compounds; some of the mechanisms used by these bacteria are conserved in mycobacteria. *Streptomyces* spp. produce diverse classes of antibiotics, e.g. macrolides, lincosamides and streptogramins. These antibiotics inhibit the bacterial ribosome by binding to the 50S ribosomal subunit. *Streptomyces* spp. are resistant to these antibiotics by expressing methyltransferases which mono- or dimethylate the adenosine residue 2058 (*Escherichia coli* notation) of the 23S rRNA, preventing the aforementioned drugs from binding to the ribosome and inhibiting translation. The *M. tuberculosis* genome encodes the methyltransferase Erm(37), a homolog of Erm methyltransferases found in many actinomycetes. However, the substrate specificity of Erm(37) differs from its homologues – Erm(37) is able to monomethylate residues 2057-2059 of the 23S rRNA, instead of only residue 2058. Monomethylation of positions 2057-2059 confers resistance to various macrolide antibiotics (Madsen *et al.*, 2005; Buriánková *et al.*, 2004).

3.4.4. Drug efflux in *M. tuberculosis*

Efflux systems are important constituents of bacterial and eukaryotic physiology. Multiple reviews have been published (Silva *et al.*, 2011; Louw *et al.*, 2009; Anthony Malinga *et al.*, 2016; Szumowski *et al.*, 2012) focusing on efflux systems in *M. tuberculosis*; the main points are briefly summarised here. Early comparative studies revealed that the genome of *M. tuberculosis* encodes a multitude of different putative efflux systems, belonging to the classes of ATP-binding cassette, major facilitator super-family, small multidrug-resistance, multidrug & toxic-compound extrusion systems and resistance-nodulation-cell division (Paulsen *et al.*, 2001).

The relevance of drug efflux for generating clinically relevant drug resistance in *M. tuberculosis* is controversial but has gained more attention in recent years. The observation that about 30 % of isoniazid (Louw *et al.*, 2009) and 3 % of rifampicin (Telenti *et al.*, 1993) resistant clinical *M. tuberculosis* isolates do not show any known resistance mutation might be explained by drug efflux. However, this unexplained resistance is potentially confounded by the fact that not all mutational targets of drug resistance are known. For certain antibiotics e.g. isoniazid, an array of different resistance mechanisms is already known (Vilchèze *et al.*, 2014). On the other hand, resistance to rifampicin is thought only to be conferred by mutations in the gene encoding one constituent of the drug target (further discussed below), making the contribution of efflux pumps to unexplained resistance phenotypes more likely.

Efflux pumps exhibit high levels of substrate promiscuity and are able to extrude a multitude of structurally unrelated compounds. Furthermore, efflux systems have been shown to be essential in *M. tuberculosis* for intracellular growth in macrophages (Lamichhane *et al.*, 2005). Mycobacterial efflux pumps are able to extrude nearly all antituberculous drugs, including streptomycin, rifampicin, isoniazid, clofazimine, bedaquiline, fluoroquinolones and ethambutol (Anthony Malinga *et al.*, 2016). Expression of efflux pumps can be viewed as a plastic trait, meaning that expression levels are modified via non-mutational processes upon changes in the environment. We can therefore say that efflux pumps are induced or upregulated when a specific environmental cue (e.g. antibiotics or the intracellular environment of a macrophage) is present. The term "overexpression" should only be used for mutants where expression levels exceed the reaction norm (Box 3.1) of the wild type strains. To our knowledge, there have not been any studies systematically investigating the reaction norm of efflux pumps in *M. tuberculosis*.

It has been demonstrated in model systems of *M. tuberculosis* that efflux pumps are induced upon infection of macrophages, which coincides with increased minimal inhibitory

concentrations for isoniazid (Adams *et al.*, 2011). A subset of the strains were resistant to higher levels of isoniazid at the peak serum concentrations (Park *et al.*, 2016). The expression of the efflux systems persists even after the mycobacterial cells have been released from the macrophages. However, in accordance with the concept of efflux pumps as a plastic trait, not all bacterial cells upregulate the expression of efflux systems (Adams *et al.*, 2011; Gupta *et al.*, 2010). Several antituberculous drugs have been demonstrated to induce the expression of efflux pump genes, but there is considerable variability between strains and no general pattern of efflux pump expression was recognizable (Gupta *et al.*, 2010). Furthermore, MDR *M. tuberculosis* isolates have been shown to constitutively express genes involved in drug efflux (Li *et al.*, 2015). Although there are reports on the upregulation of efflux systems generating minimal inhibitory concentrations slightly beyond the epidemiological cutoff (Box 3.1), the clinical relevance of efflux pumps is not clear and warrants more investigation (Adams *et al.*, 2011). The majority of drug-resistant strains harbour chromosomal mutations linked to drug resistance (further discussed below). However, there are examples of clinically relevant levels of resistance conferred by overexpression of efflux pumps. Mutations in the transcriptional repressor MmpR lead to overexpression of the multisubstrate efflux pump Mmpl5 (Milano *et al.*, 2009), which coincides with cross-resistance to clofazimine and the new antituberculous drug bedaquiline (Bloemberg *et al.*, 2015; Hartkoorn *et al.*, 2014). As Mmpl5 is also involved in isoniazid extrusion, MmpR mutants are likely also to be resistant to isoniazid (Milano *et al.*, 2009).

However, efflux systems may act as a stepping stone for the evolution of high-level resistance, as convincingly demonstrated by *in vitro* studies (Machado *et al.*, 2012), (Schmalstieg *et al.*, 2012). As efflux pumps seem to be essential for macrophage infection (Lamichhane *et al.*, 2005), efflux pump inhibitors might be used to inhibit bacterial growth and lower the MICs for certain drugs (Pule *et al.*, 2016).

3.5. Acquired drug resistance in *M. tuberculosis*

Apart from the intrinsic resistance mechanisms mentioned above, the majority of clinically relevant drug resistance in *M. tuberculosis* is conferred by chromosomal mutations. These chromosomal mutations confer drug resistance via a large array of different mechanisms and may confer different levels of resistance (Figure 3.1). The most common targets of chromosomal mutations conferring drug resistance are summarised in Table 3.1. Depending on the antibiotic in question, there may be multiple mechanisms of resistance.

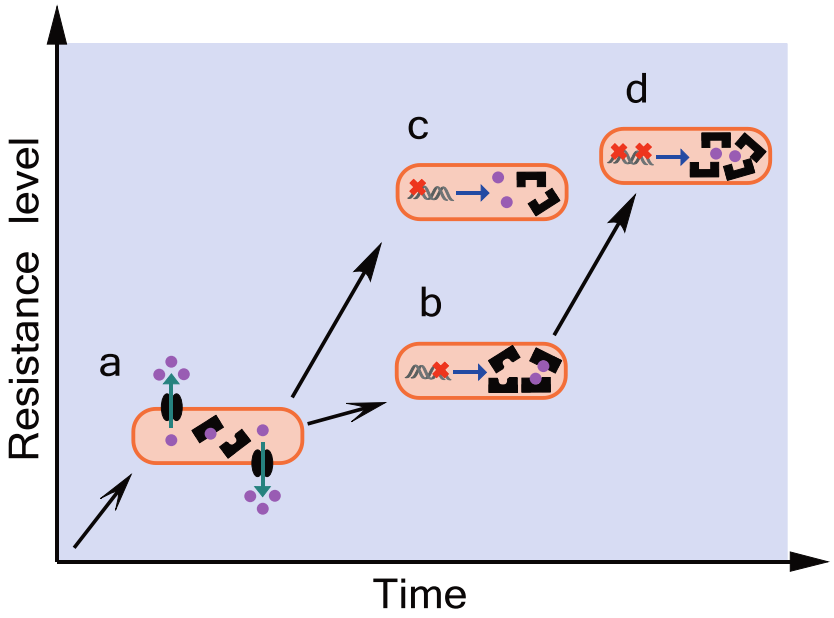

Figure 3.1.: **Levels of drug resistance conferred by different mechanisms.** The red X indicates a chromosomal mutation. **a)** Low level resistance due to induction of efflux pumps. Depiction of efflux pumps is omitted later stages for clarity **b)** Low level resistance due to target overexpression caused by chromosomal mutation. **c)** High level resistance due to drug target modification conferred by chromosomal mutation. **d)** High level resistance due to overexpression & modification of drug target conferred by independent chromosomal mutations.

3.5.1. Drug target alteration

The most common mechanism of drug resistance in *M. tuberculosis* is drug target alteration. Interactions of drug and drug target moieties are highly specific. Changes in the drug-drug target interaction sites may reduce or completely abolish drug binding and therefore confer resistance to the drug in question. Nonsynonymous mutations in drug target encoding gene(s) (Table 3.1) or nucleotide substitutions in the operon encoding the ribosomal RNA are frequently observed to confer drug resistance in *M. tuberculosis* as in the case of resistance against rifamycins, isoniazid, fluoroquinolones, aminoglycosides, cyclic peptides, para-aminosalicylic acid and oxazolidinones (e.g. linezolid). For example, mutations in the active site of the DNA-dependent RNA polymerase, corresponding to the 81 bp region known as the Rifampicin Resistance Determining Region (RRDR), confer resistance to rifampicin, by decreasing the affinity of rifampicin for target (Campbell *et al.*, 2001). Antibiotics target essential cellular functions and the drug targets perform-

ing these functions are highly conserved. The highly conserved nature of the drug targets limits the mutational target size (Box 3.1) as the resistance mutation has to accomplish two things: first it has to prevent the antibiotic from inhibiting the target and secondly it must ensure that the essential function of the drug target can still be performed. In many but not all cases this leads to a reduction in the bacterial cell's fitness in absence of the drug.

3.5.2. Abrogation of prodrug activation

Several antimycobacterial drugs are prodrugs, and abrogation of the drug activating mechanisms leads to resistance as in the case of the first-line drugs isoniazid and pyrazinamide, the second-line drugs ethionamide and para-aminosalicylic acid, as well as the two new nitroimidazole drug candidates delamanid and pretonamid. In certain cases, the prodrug activating enzyme is not essential for mycobacterial growth and survival (e.g. *pncA*/*ddn* – Table 3.1). The target size for drug resistance-conferring chromosomal mutations is therefore large – many point mutations, insertions/deletions, insertion of mobile genetic elements etc. will cause disruption of the prodrug-activating gene product without compromising bacterial survival. Furthermore, mutations in the promoter of the gene might lead to lower transcript and therefore lower levels of the enzyme activating the prodrug. Lower levels of the prodrug activating enzyme will then in turn lead to higher minimal inhibitory concentrations for the drug in question. In pyrazinamide resistant *M. tuberculosis* strains, we observe a wide array of different mutations in the gene *pncA* which encodes the enzyme metabolizing pyrazinamide to its active form pyrazinoic acid. The mutational target size for delamanid/pretomanid resistance is considerably larger as multiple enzymes and cofactors are involved in the metabolism of prodrugs to their active forms. This suggests that resistance to the latter two drugs may evolve swiftly due to the large mutational target size (Box 3.1). On the other hand, the gene *katG*, encoding a catalase/peroxidase involved in the activation of isoniazid is required for robust replication of *M. tuberculosis* in macrophages (Manca *et al.*, 1999). The mutational target size for isoniazid resistance is small, compared to pyrazinamide or delamanid/pretomanid. The resistance conferring mutation in *katG* must retain the (*in vivo*) essential function of the enzyme (catalase/peroxidase – detoxification) as well as prevent the activation of isoniazid. Most clinical *M. tuberculosis* isolates harbour the point mutation KatG S315T which retains most catalase/peroxidase functions as well as conferring high level isoniazid resistance (Pym *et al.*, 2002). On the other hand, KatG is not essential for *in vitro* replication

– this greatly enlarges the mutational target size for *in vitro* resistance, as any mutation disrupting the function of KatG will lead to resistance (Bergval *et al.*, 2009).

3.5.3. Overexpression of drug targets

Overexpression of the drug target may overcome the inhibition by the drug in question due to an overabundance of the target. Mutations in transcriptional repressors or the promoter of the drug target may cause the overexpression of the drug target as in the case of isoniazid, ethambutol and cycloserine. Drug target overexpression confers low level resistance (e.g. to isoniazid or cycloserine), which usually can be overcome by increasing the dosing of drugs administered. Drugs are administered at fixed doses, often adjusted for patient weight/age. This is generally done to achieve the maximum effectiveness of the drug whilst minimizing adverse effects of administered drugs. As certain antibiotics (e.g. cycloserine (Desjardins *et al.*, 2016a)) show dramatic adverse effects, the dose given to patients is reduced as much as possible, which means there is little room for increasing the drug doses to overcome resistance due to drug target overexpression. Overexpression of drug targets may serve as a stepping stone to high level resistance, which is conferred either by drug target alteration or abrogation of prodrug activation (Figure 3.1).

3.6. Evolution of drug resistance in *M. tuberculosis*

The evolution of drug-resistant *M. tuberculosis* variants has generally been attributed to inadequate implementation of control measures, interrupted drug supply, low quality drugs, and patient non-adherence. However, it is increasingly evident that these factors alone are insufficient to explain the evolution of drug resistance in TB, as resistant *M. tuberculosis* strains evolve in well-functioning health systems and under strict treatment adherence (Calver *et al.*, 2010; Caminero, 2008). Pathogen and host determinants are increasingly recognized to influence the evolution of drug resistance. For instance, the emerging field of pharmacogenomics has demonstrated that the current dosage regimen for TB treatment fails to generate sterilizing concentrations of certain antituberculous drugs in all patients and may contribute to treatment failure, as well as facilitate the evolution of drug resistance (Gumbo, 2010; Swaminathan *et al.*, 2012). Furthermore, recent studies have demonstrated that there is considerable variability in drug penetration into TB lesions, generating spatial and temporal variation in drug concentrations within the infected lung (Prideaux *et al.*, 2015). It is well-established that sub-inhibitory drug concentrations facilitate the evolution of drug resistance (Andersson *et al.*, 2014; Gillespie

et al., 2005; Gullberg *et al.*, 2011), and heterogeneity between and within patients may mean that some TB cases are being unwittingly exposed to sub-inhibitory treatment regimes.

Although the genus *Mycobacterium* in general and *M. tuberculosis* in particular show low genetic diversity compared to other bacteria (Achtman, 2008), the standing genetic diversity (Box 3.1) exhibited by *M. tuberculosis* translates into phenotypic diversity. There are seven extant *M. tuberculosis* lineages, which demonstrate specific phylogeographic patterns. Lineages 5 and 6 are restricted to West Africa and lineage 7 to the Horn of Africa. Lineage 1 is found along the rim of the Indian Ocean and lineage 3 is predominantly found in East Africa and South Asia. Lineage 4 and to a lesser extent lineage 2 are globally distributed (Borrell *et al.*, 2011; Coscolla *et al.*, 2014). Lineage 2 and lineage 4 have been frequently associated with drug resistance (Cohen *et al.*, 2015; Fenner *et al.*, 2012).

To better understand the *de novo* evolution of drug resistance, it is helpful to separate the different processes involved. The rate of evolution of resistance (Ford *et al.*, 2013) and the effect of drug resistance on bacterial life history traits like growth rate/yield may differ as a function of the strain genetic background. Drug resistance is often associated with reduction in bacterial fitness (Box 3.1) in the absence of the drug. This reduction is, however, not universal and can vary as a function of the genetic background (Gagneux *et al.*, 2006b). The fitness cost of drug resistance may be ameliorated by secondary, so-called compensatory mutations, which do not contribute to resistance on their own (Casali *et al.*, 2012; Comas *et al.*, 2012; Hughes *et al.*, 2013; Vos *et al.*, 2013). Furthermore, several drug resistance conferring mutations present in a single strain might interact epistatically (Box 3.1) and influence bacterial fitness (Borrell *et al.*, 2013; Gagneux *et al.*, 2006b). This means that the strain genetic background, compensatory mutations and the presence of multiple resistance mutations may interact to influence the fitness of drug-resistant *M. tuberculosis* strains (Figure 3.3).

The spread and maintenance of resistant variants in the population not only depends on the effect of drug resistance on the bacterial life history traits, but is also strongly dependent on the pathogen's population structure and effective population size, as well as genetic drift (Box 3.1). In the case of *M. tuberculosis*, genetic drift is thought to strongly influence the genetic diversity of the organism, as the effective population size is believed to be small and subject to large bottlenecks during patient-to-patient transmission (Box 3.1) (Hershberg *et al.*, 2008).

Table 3.1.: **List of the most common targets of chromosomal mutation conferring drug resistance in *M. tuberculosis*.**

Antibiotic	**Target gene**	**Resistance mechanism**	**Reference**
Rifampicin	*rpoB*	Drug target alteration	(Telenti *et al.*, 1993)
Isoniazid	*katG*	Abrogated prodrug activation	(Heym *et al.*, 1995)
	inhA	Drug target alteration	(Banerjee *et al.*, 1994; Hazbón *et al.*, 2006; Morlock *et al.*, 2003)
	inhA promotor	Drug target overexpression	(Hazbón *et al.*, 2006; Morlock *et al.*, 2003)
Ethambutol	*embB*	Drug target alteration	(Sreevatsan *et al.*, 1997)
Pyrazinamide	*pncA*	Abrogated prodrug activation	(Konno *et al.*, 1967; Scorpio *et al.*, 1996)
Ethionamide	*inhA*	Drug target alteration	(Banerjee *et al.*, 1994; Hazbón *et al.*, 2006; Morlock *et al.*, 2003)
	inhA promotor	Drug target overexpression	(Hazbón *et al.*, 2006; Morlock *et al.*, 2003)
	ethA	Abrogated prodrug activation	(Morlock *et al.*, 2003)
Fluoroquinolones	*gyrA/B*	Drug target alteration	(Malik *et al.*, 2012; Takiff *et al.*, 1994; Xu *et al.*, 1996)
Streptomycin	*rrs*	Drug target alteration	(Maus *et al.*, 2005; Meier *et al.*, 1994)
	rpsL	Drug target alteration	(Meier *et al.*, 1994; Nair *et al.*, 1993)
Amikacin	*rrs*	Drug target alteration	(Alangaden *et al.*, 1998; Maus *et al.*, 2005)
Kanamycin A	*rrs*	Drug target alteration	(Alangaden *et al.*, 1998)
	eis promotor	Overexpression of drug inactivating enzyme	(Kambli *et al.*, 2016)
Capreomycin	*rrs*	Drug target alteration	(Maus *et al.*, 2005)
	tlyA	Abrogation of drug target methylation	(Maus *et al.*, 2005; Monshupanee *et al.*, 2012)
P-aminosalicylic acid	*thyA*	Drug target bypassing	(Minato *et al.*, 2015; Zhao *et al.*, 2014)
	folC	Abrogation of prodrug activation	(Minato *et al.*, 2015; Zhao *et al.*, 2014)
Cycloserine	*ald*	Overabundance of drug target substrate	(Desjardins *et al.*, 2016a)
	alr	Drug target alteration	(Desjardins *et al.*, 2016a)
	alr promotor	Drug target overexpression	(Desjardins *et al.*, 2016a)
Bedaquiline	*atpE*	Drug target alteration	(Huitric *et al.*, 2010)
	Promotor/mmpR	Overexpression of efflux pump Mmpl5	(Bloemberg *et al.*, 2015; Hartkoorn *et al.*, 2014)
Linezolid	*rplC*	Drug target alteration	(Beckert *et al.*, 2012)
	rrl	Drug target alteration	(Hillemann *et al.*, 2008)
Delamanid Pretonamid	*ddn*	Abrogation of prodrug activation	(Bloemberg *et al.*, 2015; Haver *et al.*, 2015; Manjunatha *et al.*, 2006)
	fgd1	Abrogation of prodrug activation	(Bloemberg *et al.*, 2015; Haver *et al.*, 2015; Manjunatha *et al.*, 2006)
	fbiA/B/C	Abrogation of prodrug activation	(Bloemberg *et al.*, 2015; Haver *et al.*, 2015; Manjunatha *et al.*, 2006)
Clofazimine	*Promotor/mmpR*	Overexpression of efflux pump Mmpl5	(Bloemberg *et al.*, 2015; Hartkoorn *et al.*, 2014)

3.7. *De novo* evolution of drug resistance

There are three important factors influencing the *de novo* evolution of drug resistance: the population size, as it relates to the number of binary fission events the population has undergone, the mutation rate, and the mutational target size (Box 3.1). Together, these factors determine the rate of resistance acquisition. The strain genetic background may influence any of these parameters. It is, however, not trivial to determine the contribution of each of the factors, as they are either difficult to study *in vivo* or are not independent of each other.

3.7.1. Population size

In a series of iconic experiments, Luria and Delbrück demonstrated that, for simple traits (e.g. most bacteriophage/antibiotic resistance), bacterial populations which undergo a sufficient number of doubling events inevitably harbour resistant variants, following what is now called a Luria-Delbrück distribution (Luria *et al.*, 1943). The larger the population, the more cell division events the population experienced, and therefore the larger probability for a drug resistance mutation to arise. Furthermore, if a resistance conferring mutation evolves early during population expansion, the vast majority of the population will be resistant to a given drug even before treatment onset. We unfortunately do not have good estimates of the number of *M. tuberculosis* cells present in human lungs during infection. The number of bacteria present in a single lesion is estimated to be in the order of 10^8 bacterial cells per lesion (Shimao, 1987), although it is not clear what the basis of this estimate is. The best estimates of viable *M. tuberculosis* cells were obtained from the lungs of cynomolgus macaques (*Macaca fasicularis*) infected with *M. tuberculosis*. The disease presentation in macaques resembles that observed in humans. The study demonstrated that a macaque lung contains anything between $\approx 10^5$ and 5 x 10^8 cells, depending on the individual macaque (Lin *et al.*, 2009). However, most monkeys did not develop cavitary disease. It is believed that *M. tuberculosis* may reach very high cell densities when replicating on the interior surface of open (i.e. with access to the airways) lung cavities, but the exact cell numbers have, to our knowledge, never been assessed in humans. There is evidence that drug resistant *M. tuberculosis* variants are predominantly found where cell densities are high, i.e. at the interior of cavitating granulomas (Kaplan *et al.*, 2003a). It is exceedingly difficult to obtain reliable estimates of bacterial cell numbers in the lungs of TB patients. Recent studies have highlighted that the disease presentation in the lung is far more dynamic as previously believed (reviewed in Lenaerts, Barry and Dartois 2015). The number of granuloma cannot directly be used to estimate the number

of bacterial cells present in a patient lung, as there is a large diversity of microenvironments within granuloma and not all environments allow bacterial replication. However, if we solely focus on cavitating granuloma, we might be able to use disease severity, i.e. the number and extent of cavitations as a very rough proxy for population size. The fewer cavities present in a patient's lung, the smaller the population size of tubercle bacilli. There are reports of differences between lineages in terms of disease severity (reviewed in (Coscolla *et al.*, 2010; Coscolla *et al.*, 2014)). Lineages 2 and 4, especially the so-called "Beijing" sublineage of lineage 2, have been associated with more severe disease presentation compared to other lineages. The *M. tuberculosis* lineages 2 and 4 are also associated with drug resistance (Casali *et al.*, 2014; Cohen *et al.*, 2015; Ford *et al.*, 2013; Merker *et al.*, 2015; Mokrousov *et al.*, 2012; Niemann *et al.*, 2010; Pardini *et al.*, 2009). Differences in bacterial population sizes between lineages might therefore contribute to the differential association of specific lineages with drug resistance.

In general, we can say that the probability of evolving resistance is dependent on the number of binary fission events. *M. tuberculosis* is an intracellular pathogen and its primary niche is the macrophage. Although *M. tuberculosis* is able to survive and replicate inside macrophages, the tubercle bacteria seem to be inhibited in growth when the bacterial numbers are small (Welin *et al.*, 2011). It is not clear if this is due to killing or growth inhibition by the macrophage. However, there is evidence that the innate immune system is able to clear *M. tuberculosis* infections before the onset of adaptive immunity (Verrall *et al.*, 2014). It is therefore likely that a proportion of the tubercle bacilli are killed by the macrophage. This means that a *M. tuberculosis* population of a given size has likely undergone more binary fission events than expected from exponential growth, as the immune system continuously removes bacterial cells from the population. As killing of tubercle bacilli by macrophages is likely to occur stochastically, one can view this as a form of genetic drift (Box 3.1), potentially slowing the rate of drug resistance evolution. A recent study demonstrated that there are differences in the replication potential between different *M. tuberculosis* lineages – lineage 2 and 4 being proficient in replicating in macrophages, and strains belonging to lineage 3 and *Mycobacterium africanum* less so (Reiling *et al.*, 2013). However, it is not clear if the high cell densities observed in macrophages infected with *M.tuberculosis* strains belonging to lineage 2 and 4 are due to better survival in the macrophage or faster growth or both. In conclusion, it is not entirely clear if there is substantial variability in terms of number of binary fission events between lineages of *M.tuberculosis* The recent adaptation of a fluorescence dilution assay to *M.tuberculosis* offers the exciting possibility to assess the number of doubling events and the extent of cell death experienced by mycobacterial bacilli during macrophage in-

fection (Helaine *et al.*, 2016). Measuring growth properties of different *M.tuberculosis* lineages in an *ex vivo* system will help to clarify if there are differences between lineages concerning the number of binary fission events and therefore the rate of drug resistance evolution.

3.7.2. Mutation rates

The mutation rate is thought to be largely defined by the replication fidelity of the bacterial DNA polymerases. The basal mutation rate of *M.tuberculosis* is difficult to study as the long generation time makes mutation accumulation experiments unfeasible. A workaround is expressing the *M.tuberculosis* DnaE1 DNA polymerase heterologously in *M.smegmatis*. The generation time of *M. smegmatis* is much lower compared to *M.tuberculosis*, but it has the drawback that the effect of the genetic background (Box 3.1) on mutation rates cannot be taken into account. Using the heterologeous DnaE1 expression approach, the mutation rate has been determined to be 4.52 x 10^{-10} (2.95-7.35 x 10^{-10} 95 % confidence interval) per bp and generation, which is on the lower end of the spectrum compared to other bacteria (Rock *et al.*, 2015). Other studies have assessed the mutation rate using whole genome sequencing (Ford *et al.*, 2011). However, there are potentially many more factors influencing the mutation rate, including the intracellular environment of macrophages which are rich in reactive oxygen and nitrogen species, expression of error-prone DNA polymerases (DnaE2) due to the stressful macrophage environment, enzymes involved in repair – replication – recombination, existing drug resistance mutations, exposure to UV radiation and/or desiccation during aerosolisation (McGrath *et al.*, 2014).

3.7.3. Mutational target size

Given a mutation rate, the number of potential sites which may be mutated to confer drug resistance is an important factor involved in determining the rate of drug resistance evolution. As mentioned earlier, the mutational target size varies depending on resistance mechanism (Box 3.1). The mutational target size for prodrugs activated by nonessential enzymes is much larger compared to the target sizes for mutational drug target alteration. The mutational target size for drug target alteration is in the range of 81 bp in the gene *rpoB* for rifampicin resistance (Musser, 1995) and 117 bp in the gene *gyrA* for fluoroquinolone resistance (Maruri *et al.*, 2012). The majority of resistance-conferring mutations are found in these regions in clinico (Sandgren *et al.*, 2009). Furthermore, the

number of mutations conferring drug resistance is dependent on the resistance level in question. Drug resistance mutations differ in the level of resistance they confer. There is an inverse relationship between the drug resistance level that mutations confer and the mutational target size – the higher the selective concentration, the smaller the target size, i.e. there are fewer mutations conferring high-level than low level resistance (Ford *et al.*, 2013). Drug resistance mutations interact epistatically with the genetic background of a strain – a given drug resistance mutation may have different trait effects (Box 3.1) in different genetic backgrounds. This may lead to a reduction in the mutational target size in the broader sense when the interaction of a resistance mutation with the strain genetic background results in a detrimental trait effect in absence of the selective agent.

3.7.4. Rate of drug resistance acquisition

The rate of drug resistance acquisition is defined by the mutation rate and the mutational target size in the broader sense and may be calculated by Luria-Delbrück fluctuation assays (Luria *et al.*, 1943). In the past, there has been considerable debate on the influence of the genetic background on the rate of resistance acquisition. As mentioned previously, lineage 2 and lineage 4 have been disproportionately associated with drug resistance. It was found previously that *M.tuberculosis* strains belonging to lineage 2 carry mutations in genes involved in DNA replication, repair and recombination (Mestre *et al.*, 2011), potentially elevating mutation rates. However, this would imply that lineage 2 in general should demonstrate greater average genetic diversity compared to other lineages, which does not appear to be the case (Coscolla *et al.*, 2014). Multiple studies have focused on the rate of resistance acquisition in *M.tuberculosis*. The rate of resistance acquisition is influenced by the basal mutation rate and the mutational target size in the broader sense. Some studies report differences in resistance acquisition rates between lineages, where strains from the Beijing sublineage of lineage 2 show higher rates of resistance acquisition compared to other lineages (Ford *et al.*, 2013; Steenwinkel *et al.*, 2012), others reported similar rates between lineages (Werngren *et al.*, 2003). If there are differences in the rate of resistance acquisition between lineages, this could indicate that the mutational target size in the broader sense is different between lineages. Even if the results are discrepant, we are still missing a large part of the picture. Due to the labour-intensive nature of conducting Luria-Delbrück fluctuation assays with *M. tuberculosis*, all studies have so far focused on the rate of acquisition of resistance to rifampicin; there are no data available on differences in drug resistance acquisition rates between lineages for other drugs commonly in use to treat TB.

3.7.5. *De novo* evolution of drug resistance & treatment with drug combinations

TB cases caused by drug-susceptible *M.tuberculosis* strains are treated with a combination of 4 drugs to limit the evolution of drug resistance, as a strain would need to acquire at least 4 independent resistance mutations in order to achieve high level resistance, which is theoretically unlikely to happen if the *M.tuberculosis* bacilli carry no pre-existing resistance mutations and are exposed to all drugs at the same time. However, in some patients the tubercle bacilli evolve drug resistance despite strict treatment adherence (Calver *et al.*, 2010). This may be caused by functional monotherapy and/or subinhibitory drug concentrations as discussed above, known to facilitate the evolution of drug resistance (Andersson *et al.*, 2014; Gillespie *et al.*, 2005). Suboptimal drug concentrations can arise due to the differential potential of certain drugs to penetrate bacteria containing lesions in the human lung. For instance, pyrazinamide and rifampicin have been demonstrated to be able to diffuse and accumulate in the hypoxic and acidic granuloma, whereas the other first-line drugs do not accumulate in these structures (Prideaux *et al.*, 2015). These processes are likely to generate spatial and temporal variation in drug concentrations, facilitating the evolution of drug resistance (Moreno-Gamez *et al.*, 2015). Drug penetration into TB lesions is dependent on multiple factors including lipophilicity and solubility. Detailed knowledge of the chemical properties needed for effective distribution throughout all *M. tuberculosis* containing lesions will aid the design of novel, more effective antituberculous drug regimens (Dartois, 2014).

3.7.6. Drug resistance levels & the strain genetic background

Apart from the rate of resistance evolution, the genetic background may influence the level of resistance conferred by a drug resistance mutation. In the case of isoniazid resistance, it has been shown that the level of resistance conferred by different mutations varies with the genetic background of *M.tuberculosis*, whereby the isoniazid resistance mutation KatG S315T conferred lower levels of resistance in strains belonging to lineage 1 compared to lineage 2, 3, and 4 (Fenner *et al.*, 2012). This phenomenon could also contribute to different mutational target sizes in the broader sense in different genetic backgrounds (Ford *et al.*, 2013). Depending on the concentration of the drug, fewer resistance mutations might be available to certain lineages, which in turn would lower the rate of resistance evolution.

3.7.7. Fitness of drug-resistant *M. tuberculosis* strains

Drug resistance was long believed to be universally associated with a reduction in the drug-resistant organism's fitness in the absence of antibiotics (Box 3.1). In the case of mycobacteria this dogma was established by early studies performed with isoniazid resistant *M.tuberculosis* variants. These studies demonstrated a marked reduction in virulence of certain, but not all, isoniazid resistant strains in guinea pigs and mice (Barnett *et al.*, 1953; Middlebrook *et al.*, 1953). The observation that most drug resistance is associated with a reduction in fitness is further corroborated by studies performed in other pathogenic microorganisms (Melnyk *et al.*, 2015). Early mathematical models predicted that MDR-TB would remain a localized public health problem (Dye *et al.*, 2001; Dye *et al.*, 2002). Hence, initially, the WHO recommended focusing on Directly Observed Therapy Short Course (DOTS) for patients with drug-susceptible TB and against treating patients infected with MDR *M.tuberculosis* strains, as treating these patients is exceedingly expensive and the MDR strains were not believed to transmit enough to establish a sustainable infection chain (Espinal *et al.*, 2005).

Mathematical models predict that the probability of a drug-resistant strain spreading in the absence of antibiotic pressure is largely dependent on the resistant variant's reproductive fitness (Blower *et al.*, 2004; Cohen *et al.*, 2004; Knight *et al.*, 2015; Luciani *et al.*, 2009). The ultimate measure of fitness would be the effective reproductive number R (Box 3.1), which measures the average number of secondary cases generated per infected individual in a population of susceptible and resistant hosts. The effective reproductive number R is a high level composite measure comprising all aspects concerning transmission, including pathogen, host and environmental factors. Changes in R due to drug resistance mutations can lead to changes in allele frequencies over time. However, assessing R *in vivo* requires prospective cohort studies which are labour-intensive and expensive to conduct. Although cohort studies allow assessment of epidemiological factors associated with transmission, they often suffer from small sample sizes and are restricted to a single setting relevant for transmission (e.g. households).

To better study the impact of drug resistance on the fitness (i.e. R) of *M. tuberculosis*, it is helpful to separate the effect of drug resistance on bacterial life history traits – the trait effect (Box 3.1), from effects of drug resistance on allele frequencies in the population – the selective effect (Hall *et al.*, 2015). Trait effects of drug resistance mutations may be assessed by measuring life history traits like growth rate/yield of drug-resistant *M. tuberculosis* variants *in vitro*. Drug resistance mutations may have diverse pleiotropic effect on bacterial physiology resulting in a reduced *in vitro* growth (Figure 3.2). How-

ever, this reduction in growth yield/rates is not universal; it is strongly dependent on the mechanism of resistance, as well as the specific mutation in question. Furthermore, there is a strong effect of the strain genetic background on the fitness costs. Fitness costs may be ameliorated by secondary, so-called compensatory mutations, which do not contribute to drug resistance on their own. Compensatory mutations may lead to the retention of drug-resistant *M. tuberculosis* variants in the population in absence of selective pressure due to the drug. Epistatic interactions between the genetic background and the resistance conferring mutations seem to be pervasive and an important determinant in shaping the population biology of drug-resistant *M. tuberculosis* variants. Measuring *in vitro* fitness of drug resistance mutations as a proxy for *in vivo* fitness has been criticized for not capturing the complex dynamics involving nutrient limitation, activity of the immune system, host genetics and comorbidities (Björkman *et al.*, 2000). However, the frequency of resistance alleles in the population correlates well with the *in vitro* fitness of the strains carrying mutations for rifampicin, fluoroquinolone and aminoglycoside resistance in laboratory adapted, as well as in clinical *M. tuberculosis* isolates (Borrell *et al.*, 2013; Böttger *et al.*, 2008; Böttger *et al.*, 1998; Gagneux *et al.*, 2006b; Sander *et al.*, 2002). However, all *in vitro* fitness cost assessments are based on the ability of the *M. tuberculosis* strain being able to readily grow in artificial growth media. It is well established that not all bacilli present in patient sputum will demonstrate *in vitro* growth. This is problematic as our frequency estimates for drug resistance alleles are based on culturing *M. tuberculosis* bacilli from patient sputum, isolating genetic material and subsequent sequencing. However, there is no *a priori* reason to assume that the frequencies of drug resistance mutations are different in the populations of culturable and unculturable bacilli. Culture-free metagenomic approaches will offer a less biased view on the frequencies of drug resistance mutations present in the lungs (sputa) of TB patients (Koch *et al.*, 2014).

3.7.8. *In vitro* fitness of drug-resistant *M. tuberculosis* strains

There are two methods for assessing *in vitro* fitness of resistant variants: growth rate measurements and competition assays. Growth rate assessment focuses on the replication rate of a strain during exponential growth in absence of any drug pressure. In competition assays, the resistant variant is co-cultured with the susceptible wild type strain for one or several bacterial growth cycles in absence of the drug. At the endpoint of the experiment, the ratio of resistant to susceptible variants is assessed. Both methods allow calculation of selection coefficients. The advantage of competition assays is that they take growth yield as well as growth rate into account. Selection on growth rate occurs when there is

no population structure and resources are freely available to all individuals. On the other hand, growth yield is favoured in structured populations where resources are restricted and efficiency in nutrient utilisation is favoured (Frank, 2010). Populations of *M. tuberculosis* in human lungs have been thought of as highly structured, as observed by the concurrent presence of strains carrying different resistance mutations in separate parts of the of the infected lung (Kaplan *et al.*, 2003a; Lieberman *et al.*, 2016; Post *et al.*, 2004).

Several studies have investigated the effect of drug resistance mutations on *in vitro* fitness of *M. tuberculosis* or *M. smegmatis* either using competition assays or growth rate assessments. These studies have focused on assessing the impact of rifampicin (Gagneux *et al.*, 2006b; Song *et al.*, 2014), aminoglycoside (Freihofer *et al.*, 2016; Sander *et al.*, 2002) and fluoroquinolone (Borrell *et al.*, 2013) resistance on the bacterium's fitness compared to the susceptible ancestor in absence of the drug. We do not have reliable data on isoniazid resistance as the *in vitro* resistance mechanism differs from that observed *in vivo* (Bergval *et al.*, 2009). However, there is reason to assume that there are no or very low fitness costs of the most frequent isoniazid resistance mutation (KatG S315T), as *M. tuberculosis* variants carrying this mutation transmit efficiently and are virulent in mice (Doorn *et al.*, 2006; Gagneux *et al.*, 2006a; Pym *et al.*, 2002; Soolingen *et al.*, 2000). For the same reasons we do not have reliable data on the fitness of clinical MDR/XDR strains. We are restricted to measuring the fitness of such strains relative to each other – rather than to their cognate drug-susceptible ancestor (Naidoo *et al.*, 2014; Spies *et al.*, 2013).

Most rifampicin resistance mutations carry a significant fitness cost in *M. tuberculosis* (Billington *et al.*, 1999; Gagneux *et al.*, 2006c; Koch *et al.*, 2014; Kaplan *et al.*, 2003b) (Figure 3.2). However, there are some mutations (e.g. RpoB S450L – *M. tuberculosis* notation (Andre *et al.*, 2017)) which only cause a small defect in fitness. There seems to be some influence of the genetic background on the cost of rifampicin resistance, whereby lineage 2 experiences smaller reductions in fitness compared to lineage 4 (Gagneux *et al.*, 2006b). Interestingly, the fitness costs of rifampicin resistance is elevated under nutrient restricted conditions (Song *et al.*, 2014). Nutrient limitation is pervasive in the human host, indicating that there will be strong selection against costly drug resistance mutations. Fluoroquinolone resistance mutations show varying degrees of costs in *M. smegmatis* (Borrell *et al.*, 2013). Some mutations conferring fluoroquinolone resistance even show a fitness benefit, but this could be an artefact of the *in vitro* assessment conditions. Most aminoglycoside resistance is also costly. There are, however, also resistance mutations which do not demonstrate any discernable fitness cost (Freihofer *et al.*, 2016; Sander *et*

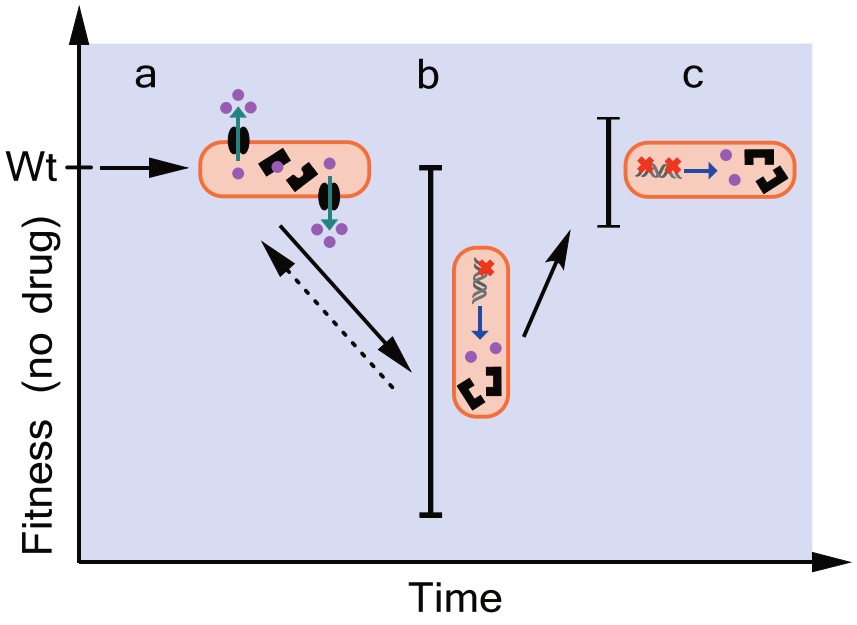

Figure 3.2.: **Fitness of drug resistant *M. tuberculosis* strains in absence of the drug**. Solid arrows indicate the trajectories of drug resistance. Dashed arrow indicates the unlikely possibility of reversion. The red X indicates a chromosomal mutation **a)**Induction of efflux pumps at no or minimal fitness cost. **b)** Large variation in fitness costs due to resistance via chromosomal mutations – ranging from very low/no to lethal. The observed variation may arise due to epistatic interactions of the strain genetic background and the resistance mutation(s). **c)** Variation in fitness levels of strains harbouring secondary compensatory mutations – ranging from comparable to wild type levels to slightly below.

al., 2002). The effect of the strain genetic background on the fitness cost of aminoglycoside resistance is yet to be determined. The *in vitro* fitness of MDR and especially XDR *M. tuberculosis* strains is variable. Some strains demonstrate very low replication rates, whilst others replicate at rates similar to drug-susceptible strains (Naidoo *et al.*, 2014; Spies *et al.*, 2013). However, it is striking that combinations of low cost resistance mutations rifampicin (RpoB S450L – *M. tuberculosis* notation) and isoniazid (KatG S315T) in MDR/XDR strains are frequently observed in clinico. TB caused by XDR *M. tuberculosis* strains is a rather new phenomenon. It is thought that most XDR *M. tuberculosis* isolates arise due to *de novo* evolution rather than transmission (Casali *et al.*, 2014; Gandhi *et al.*, 2006). The isolation of XDR *M. tuberculosis* strains with very low replication rates could reflect the fact that selection has not yet removed the most unfit XDR strains from the population. However, it is remarkable that there is a strong association of MDR/XDR strains with lineages 2 and 4 (Casali *et al.*, 2014; Cohen *et al.*, 2015; Ford *et al.*, 2013; Merker *et al.*, 2015; Mokrousov *et al.*, 2012; Niemann *et al.*, 2010; Pardini *et al.*, 2009).

This might indicate that strains belonging to these lineages experience lower fitness costs due to drug resistance. On the other hand, lineages 2 and 4 are also the most widespread. The association of these lineages with drug resistance could just be due to the large numbers of TB patients infected with strains from these lineages. To elucidate if the frequent association of lineages 2 and 4 with drug resistance is due to lower fitness costs of drug resistance experienced by these strains, we need detailed assessments of the population structures of drug-resistant and susceptible *M. tuberculosis* strains over time.

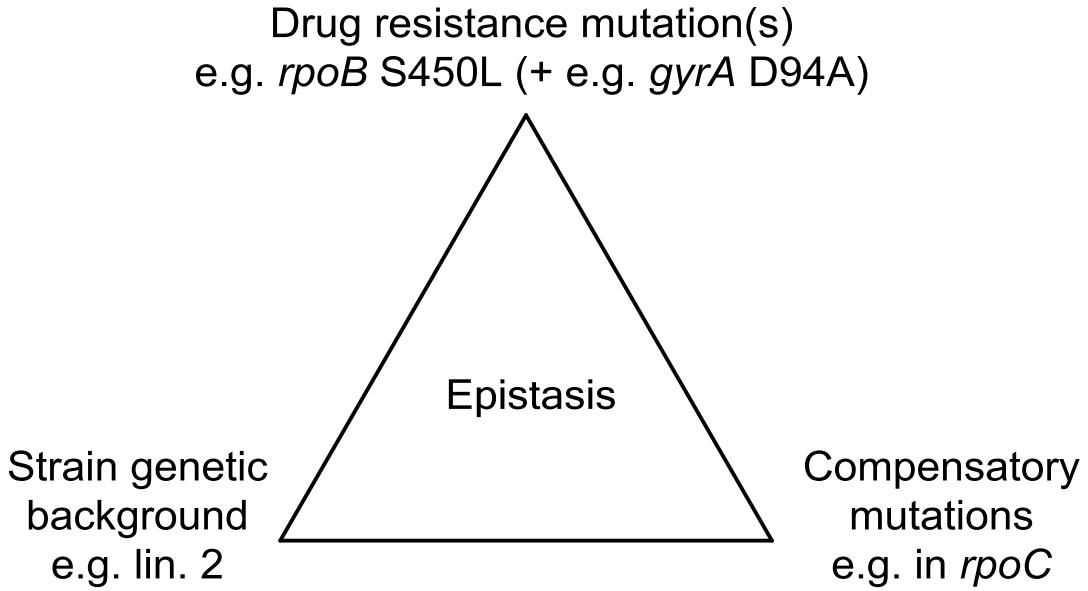

Figure 3.3.: **Pervasive epistatic interactions influencing the physiology and population biology of drug-resistant *M. tuberculosis* strains.** The effect of drug resistance mutations alone or in combination, as well as the presence of compensatory mutations affect the fitness of drug-resistant *M. tuberculosis* strains as a function of the strain genetic background.

3.7.9. Compensation of fitness costs

Fitness costs caused by drug resistance mutations may be ameliorated by secondary so-called compensatory mutations (Figure 3.2). These mutations do not contribute to drug resistance directly. Fitness cost compensation results from the epistatic interaction of the compensatory mutation with the drug resistance mutation and the genetic background of the strain (Figure 3.3, Box 3.1). Fitness cost compensation is poorly understood. Whole genome sequencing studies have demonstrated that MDR and especially XDR strains

tend to harbour a multitude of mutations (Merker *et al.*, 2015; Zhang *et al.*, 2013). A high proportion of these mutations are thought to have functional consequences. As these strains are resistant to many antibiotics, the reported mutations can be speculated to be involved in compensation of fitness costs. To date, we know of three distinct mechanisms involved in the compensation of fitness costs (Andersson *et al.*, 2010). First, secondary mutations in the genes encoding the drug target might restore or improve the mutated enzyme's function to levels comparable to the wild type enzyme (Casali *et al.*, 2012; Comas *et al.*, 2012; Hughes *et al.*, 2013). To be classified as a putative compensatory mutation, these mutations should only be found in resistant strains and never in susceptible strains. However, only fitness measurements can confirm or refute the compensatory role of the mutation. Second, overexpression of an enzyme which performs a similar function to the drug target may compensate for the diminished or abolished function of the drug target (Heym *et al.*, 1997; Sherman *et al.*, 1996). Third, non-mutational gene regulatory responses may compensate for the fitness defect inflicted by the resistance mutation (Freihofer *et al.*, 2016). However, the latter case is a form of phenotypic plasticity and the term "buffering" of fitness costs would be more appropriate than compensation. Compensatory evolution has been demonstrated to occur very frequently for rifampicin resistance (Casali *et al.*, 2012; Comas *et al.*, 2012; Hughes *et al.*, 2013). Fitness costs of rifampicin resistance due to mutations in the β'-subunit of DNA dependent RNA polymerase encoded by RpoB can be compensated by mutations in *rpoABC*. The molecular basis of fitness costs and their compensation in rifampicin-resistant *M. tuberculosis* strains is not well understood. However, studies performed with rifampicin-resistant *Pseudomonas aeruginosa* have indicated that rifampicin resistance mutations in *rpoB* reduce the transcriptional efficiency of the DNA-dependent RNA polymerase and may explain the observed fitness costs. Compensatory mutations restored the transcriptional efficiency to wild type levels in rifampicin-resistant strains (Qi *et al.*, 2014). The connection between transcriptional efficiency and *in vitro* growth rate (fitness cost) seems plausible in fast-growing bacteria like *P. aruginosa*.

However, it remains to be demonstrated that fitness costs inflicted by mutations in *rpoB* are caused by reduced transcriptional efficiency in slow-growing bacteria like *M. tuberculosis*. Interestingly, the compensatory mutations are often found at the interface between the different RNA polymerase subunits. More specifically, the compensatory mutations home to positions in the vicinity of the active site and the RNA exit tunnel (Song *et al.*, 2014). Compiling data from several studies (Bloemberg *et al.*, 2015; Casali *et al.*, 2014; Comas *et al.*, 2012; Eldholm *et al.*, 2015; Farhat *et al.*, 2013; Lanzas *et al.*, 2013; Merker *et al.*, 2015; Müller *et al.*, 2013; Song *et al.*, 2014; Vos *et al.*, 2013), ri-

fampicin resistance mutations in *rpoB* and putative compensatory mutations in *rpoABC* offer some striking insights. In *M. tuberculosis*, most compensatory mutation in *rpoA/C* (Figure 3.4 b and c) are associated with the most clinically frequent rifampicin-resistance mutation RpoB S450L (Figure 3.4 a and d). The rifampicin resistance-conferring mutation RpoB S450L consistently shows the lowest fitness cost *in vitro* (Casali *et al.*, 2014; Cohen *et al.*, 2015; Meftahi *et al.*, 2016; Vos *et al.*, 2013). There are two non-exclusive explanations for this observation. First, the low average fitness cost of RpoB S450L might offer a large mutational spectrum for compensation, i.e. there are a multitude of different mutations in *rpoABC* which compensate for the fitness cost of RpoB S450L. More costly rifampicin resistance mutations might have a more restricted mutational target for compensation. An alternative explanation would be a sequential nature of resistance and compensatory mutation acquisition. *M. tuberculosis* strains carrying RpoB S450L are able to outcompete other rifampicin resistance mutations due to its low cost. This will lead to an overabundance of *s*trains carrying the RpoB S450L mutation in the population, making it more likely for these strains to acquire compensatory mutations. Loss of function mutations in the gene *katG* are known to confer resistance to isoniazid (Table 3.1). However, the catalase/peroxidase activity of KatG is essential for efficient *in vivo* growth of *M. tuberculosis*. Besides KatG, the genome of *e*ncodes the peroxidase AhpC. Certain clinical isoniazid resistant *s*trains have been noticed to harbor mutations in the regulatory regions of the gene *ahpC*, which leads to overexpression of AhpC. However, it was noticed that overexpression of *ahpC* does not compensate for the reduction in virulence observed in *katG* deletion mutants. Overexpression of *ahpC* might therefore only partially restore the fitness costs of *katG* deletions by enhancing detoxification of detrimental reactive oxygen species (Heym *et al.*, 1997; Sherman *et al.*, 1996). Recently, a non-mutational mechanism of fitness cost compensation was described. Upregulation of *tlyA* expression in capreomycin-resistant *s*trains was found to partially compensate (buffer) the fitness cost inflicted by the mutation conferring resistance to capreomycin (Freihofer *et al.*, 2016). Resistance to capreomycin can be conferred by mutations in the ribosomal RNA (Table 3.1). The gene *tlyA* encodes a methyltransferase which methylates nucleotides of the ribosomal RNA. Methylation of the ribosomal RNA nucleotide directly adjacent to the resistance conferring mutated nucleotide seems to partially restore fitness levels (Freihofer *et al.*, 2016). Mitigation of fitness costs via non-mutational processes can be viewed as a plastic trait.

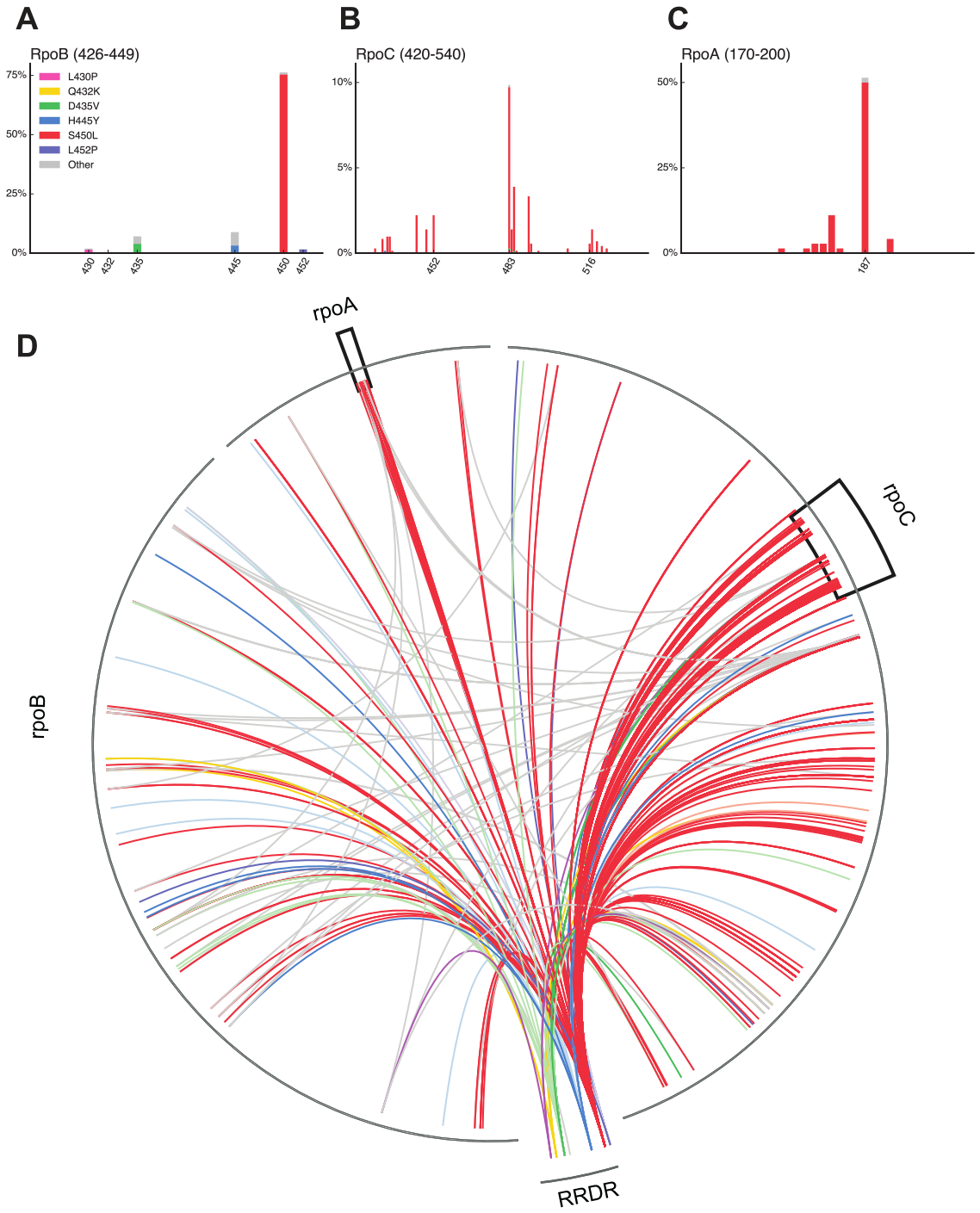

Figure 3.4.: **Summary of rifampicin resistance and fitness cost compensatory mutations in *rpoB* and *rpoA/C* respectively.** **(a)** Frequency of rifampicin resistance mutations in *rpoB*. **(b)** Frequency of putative compensatory mutations in *rpoC* affecting codons 420-540. **(c)** Frequency of putative compensatory mutations in *rpoA* affecting codons 170-200. **(d)** Association of rifampicin resistance mutations in *rpoB* with putative compensatory mutations in *rpoA/C*.

3.7.10. *In vivo* fitness of drug-resistant *s*trains

Although *in vitro* fitness measurements have provided us with valuable estimates of the fitness costs inflicted by drug resistance, the ultimate measure incorporating all steps from establishing an infection to transmission of *a*re captured by the effective reproductive number R (Box 3.1). Several studies have attempted to approximate R by conducting prospective cohort studies. These have mainly focused on the number of secondary cases generated in households of TB patients infected with drug-susceptible or drug- resistant *s*trains. These studies (reviewed in Fox *et al.* 2013)) have revealed conflicting results, with some showing considerably lower transmission rates among MDR vs. susceptible *M. tuberculosis* strains and others showing comparable transmission rates. The heterogeneity in the determined transmission rates might reflect differences in study design.

Household contact studies are expensive and labour intensive to conduct. A more convenient and powerful way of assessing the *in vivo* fitness of drug-resistant *M. tuberculosis* strains is a population-based high resolution genotyping study. The combination of whole genome sequencing together with high quality epidemiological data will allow assessment of transmission rates at high spatial and temporal resolution. Several studies have already demonstrated the value of whole genome sequencing for inferring transmission networks (Casali *et al.*, 2014; Hatherell *et al.*, 2016; Luo *et al.*, 2014; Walker *et al.*, 2013; Yang *et al.*, 2016).

3.7.11. Epistatic interactions between drug resistance mutations

TB infections are always treated with a combination of drugs. There is therefore strong selection for *M. tuberculosis* strains which are concomitantly resistant to several drugs used for treatment. The effect of multiple drug resistance mutations on the organism's phenotype is dependent on the epistatic interactions of the different resistance mutations with each other and with the genetic background of the strain. Certain combinations of resistance mutations may therefore be especially favoured if they do not confer a fitness deficit. Due to the vast number of different combinations of different drug resistance mutations, it is difficult to assess the contribution of different combinations of resistance mutations to the phenotype experimentally, especially in *M. tuberculosis*. However, assessing allele frequencies from clinical isolates may inform us about successful vs. detrimental combinations of drug resistance alleles.

Several drugs target the information pathway from DNA to mRNA to proteins by inhibiting key enzymes involved in these processes. Mutations in these enzymes alone or in combination have pleiotropic effects on the fitness of *M. tuberculosis*. Infections with MDR *M. tuberculosis* strains are treated with drug regimens containing fluoroquinolones. Per definition, these strains are resistant to isoniazid and rifampicin. Fluoroquinolone resistance mutations will therefore mostly evolve in rifampicin-resistant *M. tuberculosis* strains. Rifampicin targets the DNA dependent RNA polymerase and fluoroquinolones target the DNA-gyrase and these two enzymes work in concert. Studies performed with *M. smegmatis* demonstrated that depending on the combination of rifampicin and fluoroquinolone resistance mutation, the effects on the *in vitro* growth rate range from detrimental to even beneficial. This goes so far, that fitness cost of a resistance mutation is fully compensated by the presence of a second resistance mutation. In certain cases, the double-resistant strain had an even faster growth rate than the drug-susceptible wild type. Combinations of the most beneficial resistance alleles are also the most frequent in clinical *M. tuberculosis* isolates (Borrell *et al.*, 2013). *In vitro* studies have demonstrated that pre-existing isoniazid resistance mutations influence which rifampicin resistance mutations are subsequently acquired. This is probably due to a reduction in the mutational target size in the broader sense (Box 3.1) due to epistatic interactions between the two resistance conferring mutations, resulting in a dramatic reduction in fitness of specific combinations (Bergval *et al.*, 2012). Streptomycin inhibits the bacterial ribosome and resistance is often conferred by mutations in the gene *rpsL*. Streptomycin was the first active antituberculous drug in use and resistance to this drug is widespread. Household contact studies have determined that *M. tuberculosis* strains harbouring the streptomycin resistance-conferring mutation RpsL K43R together with the isoniazid resistance mutation KatG S315T generated significantly fewer secondary cases indicating an epistatic interaction between the streptomycin and isoniazid resistance mutations resulting in a fitness cost (Salvatore *et al.*, 2016). However, this is in contrast to other studies, which have found that the combination of RpsL K43R with KatG S315T does not reduce fitness of *M. tuberculosis* harbouring these mutations (Spies *et al.*, 2013). These discrepancies demonstrate the large variability in these interactions, highlighting the influence of the strain genetic background on all these interactions. In turn, drawing general conclusions from rather small datasets is not possible or at best challenging.

3.8. Population genetics of drug resistance in *M. tuberculosis*

The forces of evolution may determine the fate of drug-resistant *M. tuberculosis* variants at multiple stages during infection. During transmission, *M. tuberculosis* populations undergo stark bottlenecks (Box 3.1) as new infections are thought to be established only by a few bacilli (Jacobs, 1941). This means that in most cases, the bacterial populations in the lungs of TB patients will demonstrate only very low levels of genetic diversity. Only about 12 % of all infections with *M. tuberculosis* will result in active disease within the life time of a patient (Dheda *et al.*, 2016). After the formation of the primary complex, *M. tuberculosis* bacilli are disseminated throughout the lung where they form new foci. The bacteria present in these foci are thought to form isolated subpopulations with little migration between the populations, at least in the early stages of the infection. Structured populations are thought to select for growth yield optimisation, as they do not have to compete for resources with other, faster growing variants due to the private nature of the resources (Frank, 2010). Indeed, sampling bacteria from different lesions revealed the presence of distinct subpopulations throughout the lungs of TB patients (Lieberman *et al.*, 2016). However, sampling *M. tuberculosis* bacteria from patient sputum during treatment at different time points revealed a dynamic picture of the bacterial population. Multiple different resistance alleles e.g. conferring resistance to fluoroquinolones were present concomitantly – a phenomenon known as heteroresistance. Furthermore, certain drug resistance alleles were increasing over time whilst others were disappearing. Moreover, there was selection for strains harbouring less costly resistance and/or known compensatory mutations (Bernard *et al.*, 2016; Bloemberg *et al.*, 2015; Eldholm *et al.*, 2014; Merker *et al.*, 2013). Selection for *M. tuberculosis* variants with higher growth rates might indicate that resources available in a patient's lung are not private and that the bacterial populations are less structured than we had anticipated. Indeed, studies performed in cynomolgus macaques and TB patients (Coleman *et al.*, 2014; Lin *et al.*, 2013) demonstrate a much more dynamic picture. During the course of *M. tuberculosis* infection, existing foci have been observed to be eliminated by the immune system, whereas new foci seem to appear, without any discernable pattern. The immune system is apparently able to eliminate foci and presumably kill the bacteria present in these lesions. The removal of bacteria from these lesions by the immune system can be viewed as a form of genetic drift. This has implications for the *de novo* evolution of drug resistance. Populations of *de novo* evolved resistant *M. tuberculosis* variants inevitably start out with low numbers. Populations with small effective population sizes are vulnerable to extinction by genetic

drift. This means that the action of the immune system might slow the *de novo* evolution of drug resistance, by stochastically removing resistant variants from the population. Host directed therapies (Zumla *et al.*, 2015) boosting the action of the immune system in clearing TB infections will inevitably also reduce the rate at which drug-resistant variants evolve *de novo*.

3.9. Conclusion and outlook

The intrinsic resistance of *M. tuberculosis* against many classes of antibiotics, the ever rising numbers of drug-resistant strains, as well as the scarcity of novel antituberculous compounds is threatening the progress in containing the disease. The remarkable capability of *M. tuberculosis* to evolve drug resistance against all efficacious drugs at low or no cost underlines the necessity of a multi-pronged strategy to reduce the incidence of both drug-susceptible and drug-resistant *M. tuberculosis* variants. Aside from improving the quality of public health systems in resource-limited settings, novel, cost-effective point-of-care diagnostic tools, drugs, as well as an effective vaccine are urgently needed. Furthermore, the influence of the strain genetic background on virtually all aspects of drug resistance evolution highlights that it is not sufficient to focus research on laboratory strains. It is crucial to include *M. tuberculosis* strains which ideally represent the whole phylogenetic space of the species in order to assess the impact of a novel drug, or for that matter, any planned intervention on the population biology of this extraordinarily successful pathogen.

Box 3.1 Commonly used terms in evolutionary biology

Bottleneck

A bottleneck describes the stark reduction of the population size and therefore the genetic diversity of the population due to random sampling, i.e. genetic drift. In the context of *M. tuberculosis* infections, only a subset of the whole population of bacteria present in the lungs of a patient will gain access to the airways and may be aerosolized. Furthermore, presumably only a subset of that aerosolized population will survive the harsh environmental conditions found outside of the host, and is able to infect a new host – if one is present to inhale the aerosols.

Effective population size

In the context of bacterial (haploid), non-recombining, obligate pathogens like *M. tuberculosis*, the effective population size describes the proportion of the population e.g. in a patient, which has the possibility to transmit to a new host. *M. tuberculosis* is able to establish infections in nearly all tissues – however, all of them are "evolutionary dead ends", pulmonary (or in rare cases laryngeal) infections being the only exception. Consider the case of a patient with pulmonary and extra-pulmonary TB. The effective population size could be defined as the total number of *M. tuberculosis* cells present in a patient minus the number of cells in the extra-pulmonary location, ergo the pulmonary (laryngeal) *M. tuberculosis* population – as only this population will be able to transmit to a new host. Furthermore, transmission of *M. tuberculosis* requires substantial lung damage that "allows" access to the bronchi, which is dependent on the location of the foci within the lung, ergo not all *M. tuberculosis* cells in the lung will have the potential to transmit, reducing the effective population size.

Epidemiological cutoff

The highest minimal inhibitory concentration of a drug observed in a wild type strain. This is related to the reaction norm – see below.

Epistasis

Epistasis describes the phenomenon where the interaction of two or more genes/alleles produce an effect on the phenotype (e.g. on fitness) which is unequal to the sum each gene's/allele's effect on their own.

Fitness

The fitness of an organism can be defined as the ability to survive and reproduce in a given environment. In the case of an obligate parasite like *M. tuberculosis* this entails establishing an infection in the human host, replicating and transmitting to a new host. Fitness can be parameterized by the effective reproductive number R, which quantifies how many secondary cases are produced on average by a single infected individual in a population of susceptible and resistant hosts. However, R is notoriously difficult to assess *in vivo*. In absence of a better measure, fitness of bacteria is often approximated by measuring life history traits e.g. *in vitro* growth rates and/or growth yield in artificial growth media. However, there is some merit in measuring *in vitro* growth rates/yields e.g. of resistant *M. tuberculosis* variants as their growth rate/yields are often correlated with *in vivo* frequency of these variants.

Fitness cost

A fitness cost describes the reduction in the number of offspring produced by a genotype in a given environment caused e.g. by a drug resistance mutation, compared to a drug-susceptible variant. Drug resistance has long been assumed to be generally associated with a fitness cost in absence of the drug (i.e. a reduction in R, the number of secondary cases generated). Using *in vitro* methods, a fitness cost would manifest itself as a reduction in growth rate or yield.

Genetic background

The genetic background describes the genetic diversity present in a strain's genome resulting from mutations/insertions/deletions/rearrangements etc. when compared to other strains. The genetic background is an important factor influencing epistatic interactions.

Genetic drift

Genetic drift is an important mechanism of evolution whereby the genetic diversity of a population is reduced by random sampling error. The sampling error results from the differential probability of an organism to survive and reproduce based on chance. The effect of genetic drift on allele frequencies is strongest in populations with a small effective population sizes. Genetic drift can lead to the stochastic fixation or loss of alleles in a given population.

Natural selection

Natural selection is one of the major mechanisms of evolution and refers to the differential survival of organisms based on their phenotype. The phenotype results from the interaction of the organism's genotype with the environment. The differential survival results from competition (e.g. for resources) among organisms and is dependent on the degree of adaptation of the organism to its environment. The effect of natural selection is strongest in large populations and leads to changes in allele frequencies over time.

Mutational target size

The mutational target size in the narrow sense describes the number of different mutations which may confer resistance to a certain drug. Depending on the resistance mechanism, the mutational target size can vary greatly. The mutational target size in the broader sense describes the number of different mutations which confer resistance to a given drug and which do not result in lethality due to epistatic interactions with the strain genetic background.

Reaction norm

The phenotype of an organism results from the interaction of the organism's genotype with the environment. The range of different phenotypes of a given genotype in different environments is called the reaction norm (Griffiths *et al.*, 2005).

Standing genetic diversity

The term standing genetic diversity describes the presence of multiple alleles at a locus which are segregating in the population. In other words, the sum of all genetic backgrounds (see above) constitutes the standing genetic diversity of a species.

Trait effect

Hall *et al.* describe the trait effect as e.g. the effect drug resistance mutations may have on life history traits like growth rate and/or yield of the strains carrying resistance mutations. The magnitude of the trait effect inflicted by drug resistance mutations is dependent on multiple factors including the genetic background, pre-existing drug resistance mutations, compensatory mutations and the environment (Hall *et al.*, 2015).

4. Whole genome sequencing for drug resistance profile prediction in *Mycobacterium tuberculosis*

Sebastian M. Gygli[1,2,*], Peter M. Keller[3,12*], Marie Ballif[5], Nicolas Blöchliger[3], Rico Hömke[3,4], Miriam Reinhard[1,2], Chloe Loiseau[1,2], Claudia Ritter[3,4], Peter Sander[3,4], Sonia Borrell[1,2], Jimena Collantes Loo[11], Anchalee Avihingsanonf[6,7], Joachim Gnokoro[8], Marcel Yotebieng[9], Matthias Egger[5,10], Sebastien Gagneux[1,2,§], Erik C. Böttger[3,4,§]

[1]Swiss Tropical and Public Health Institute, Basel, Switzerland
[2]University of Basel, Basel, Switzerland
[3]Institute of Medical Microbiology, University of Zürich, Zürich, Switzerland
[4]National Center for Mycobacteria, University of Zürich, Zürich, Switzerland
[5]Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland
[6]HIV-NAT/Thai Red Cross AIDS Research Centre, Bangkok, Thailand
[7]TB Research Unit, Department of Medicine, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand
[8]Centre de Prise en Charge de Recherche et de Formation, Yopougon, Abidjan, Côte d'Ivoire
[9]College of Public Health, Ohio State University, Columbus, Ohio, USA
[10]Centre for Infectious Disease Epidemiology and Research, University of Cape Town, Cape Town, South Africa
[11]Instituto de Medicina Tropical Alexander von Humboldt, Universidad Peruana Cayetano Heredia, Lima, Peru
[12]Institute for Infectious Diseases, University of Bern, Bern, Switzerland
[*]Authors contributed equally
[§]Corresponding authors

One sentence summary

Whole genome sequencing of clinical *M. tuberculosis* isolates accurately predicts drug resistance profiles and may replace culture-based drug susceptibility testing in the future.

Key words

Mycobacterium tuberculosis, quantitative phenotypic drug susceptibility testing, whole genome sequencing, drug resistance, drug resistance level prediction

This chapter has been published in Antimicrobial Agents and Chemotherapy, Volume 63, Issue 4, 1st of April 2019, Pages e02175-18

4.1. Abstract

Whole genome sequencing allows rapid detection of drug-resistant *M. tuberculosis* isolates. However, high-quality data linking quantitative phenotypic drug susceptibility testing and genomic data have thus far been lacking. We determined drug resistance profiles of 176 genetically diverse clinical *M. tuberculosis* isolates from Democratic Republic of the Congo, Ivory Coast, Peru, Thailand and Switzerland by quantitative phenotypic drug susceptibility testing for 11 antituberculous drugs using the BD BACTEC MGIT 960 system and 7H10 agar dilution to generate a cross-validated phenotypic drug susceptibility testing readout. We compared drug susceptibility testing results with predicted drug resistance profiles inferred by whole genome sequencing. Both phenotypic drug susceptibility testing methods identically classified the strains into resistant/susceptible in 73-99 % of the cases, depending on the drug. Changes in minimal inhibitory concentrations were readily explained by mutations identified by whole genome sequencing. Using whole genome sequences, we were able to predict quantitative drug resistance levels for many drug resistance mutations. Predicting quantitative levels of drug resistance by whole genome sequencing was partially limited due to incompletely understood drug resistance mechanisms. The overall sensitivity and specificity of whole genome-based drug susceptibility testing were 86.8 % and 94.5 %, respectively. Despite some limitations, whole genome sequencing has the potential to infer resistance profiles without the need for time-consuming phenotypic methods.

4.2. Introduction

Timely and accurate drug susceptibility testing (DST) of *M. tuberculosis* isolates is vital to prevent the transmission of multidrug-resistant strains (MDR – resistance to rifampicin and isoniazid)(World Health Organization, 2010). However, the slow growth and stringent biosafety requirements of *M. tuberculosis* make obtaining a full DST profile by culture-based techniques a matter of weeks or months. In addition, culture-based DST is notoriously challenging for several drugs, e.g. pyrazinamide and ethionamide due to poor drug solubility in commonly used culture media (Domínguez *et al.*, 2016).

Drug resistance in *M. tuberculosis* is mainly conferred by chromosomal mutations in a few genes (Gygli *et al.*, 2017), making it possible to detect drug resistance by sequencing these genes or probing them by molecular hybridisation (Deggim-Messmer *et al.*, 2016). Several commercial tests for the detection of resistance-associated mutations are available, e.g. the GenoType MTBDRplus V2 (Hain Lifescience GmbH, Nehren, DE) (Nathavitharana *et al.*, 2016), the AID TB Resistance Line Probe Assay (AID GmbH, Strassberg, DE) (Ritter *et al.*, 2014). The World Health Organisation (WHO) endorses line probe assays for the detection of rifampicin resistance as a surrogate marker for multidrug-resistance (World Health Organization, 2016b) and the GeneXpert® system (Cepheid, Sunnyvale, CA, USA) for the detection of rifampicin and isoniazid resistant strains (WHO, 2013). These molecular tests demonstrate high sensitivities for drugs with established target(s) of resistance and for which only a few mutations are responsible for most resistance in clinico (e.g. rifampicin, isoniazid) (Deggim-Messmer *et al.*, 2016). However, molecular tests show low sensitivity for heteroresistant strains (concomitant presence of wild type (wt) and resistant or multiple different resistant variants in patient isolates), when frequencies of resistant variants drop below 5-50 % (Engström, 2016; Streicher *et al.*, 2012). Furthermore, there are no commercially available tests for many drugs currently/prospectively in use (e.g. bedaquiline, delamanid, linezolid, p-aminosalicylic acid).

The past years have seen a wealth of genomic data on drug-resistant *M. tuberculosis* become available (Coll *et al.*, 2015; Walker *et al.*, 2015). However, phenotypic DST data are lacking for most of the genetic data sets. In addition, DST data are often limited as the strains were classified as susceptible or resistant using only a single drug concentration (The CRyPTIC Consortium and the 100000 Genomes Project, 2018). There is an urgent need to link genotypic and phenotypic drug resistance readouts to obtain a better understanding of the mechanisms influencing the evolution and spread of drug resistance in *M. tuberculosis*.

WGS of clinical isolates allows for accurate identification of established-resistance-conferring chromosomal mutations (Coll *et al.*, 2015; Colman *et al.*, 2016; Shea *et al.*, 2017) and may ensure adequate treatment in days instead of months. We compared whole genome-based drug resistance profiles with two culture-based quantitative DST methods for a total of 11 drugs, including all group 1 drugs (rifampicin, rifabutin, isoniazid, ethambutol, pyrazinamide), all group two (streptomycin, kanamycin A, amikacin and capreomycin), as well as selected group 3 (moxifloxacin) and group 4 drugs (ethionamide).

4.3. Material and methods

4.3.1. *M. tuberculosis* isolates

The initial data-set consisted of 189 *M. tuberculosis* isolates. A subset of 61 strains was used to establish the phenotypic DST methodology. These 61 strains were collected by the Swiss National Center for Mycobacteria between 2004-2015, and represent a broad spectrum in geographic origin and drug resistance profiles (Bloemberg *et al.*, 2015; Springer *et al.*, 2009; Stucki *et al.*, 2016). We then applied the quantitative DST methodology to 125 prospectively collected clinical isolates from clinics participating in the International Epidemiology Databases to Evaluate AIDS (IeDEA) (Egger *et al.*, 2012) in Peru, Thailand, Ivory Coast and the Democratic Republic of the Congo. Thirteen strains had to be excluded due to failed WGS (n = 4, failed library preparation due to low DNA quality), irreproducible DST results (n = 1), no growth in the 7H10 agar dilution assay (n = 3), duplication (n = 1), mixed cultures (n = 2, cross-contamination or patient infected with multiple strains) or transmission clusters (n = 2). The final set consisted of 176 strains.

4.3.2. Phenotypic DST

MGIT 960- and 7H10 agar dilution-based phenotypic DST were performed as described previously (Springer *et al.*, 2009). Critical concentrations used for the classification of strains into resistant/susceptible aim to predict clinical outcome, i.e. treatment failure if a given strain is resistant at the critical concentration. However, critical concentrations should ideally be defined on the basis of the epidemiological cut-off (ECOFF: The highest wild type MIC observed in absence of any detectable resistance mechanism (Schön *et al.*, 2017)), treatment outcomes and pharmako-kinetic and -dynamic data. However, as

M. tuberculosis infections are treated with combination therapy, outcome data for single drugs are difficult to obtain (Ängeby *et al.*, 2012). This calls for definition of critical concentrations solely based on the ECOFF (World Health Organization, 2018b). We therefore classified strains as resistant/susceptible on the basis the ECOFF derived from our data (World Health Organization, 2014a). Table 4.1 lists the epidemiological cut-offs (ECOFF) used, Table A.4.2 the drug concentrations tested with the MGIT 960 and 7H10 agar-dilution assays and Table 4.2 the genes screened for mutations with WGS. Further details are available in the supplementary materials.

4.3.3. Data analysis

The categorical agreement between the MIC determination by MGIT 960 and 7H10 agar dilution was determined based on the ECOFF (Table 4.1). The numerical variation between the two methods was quantified as the geometric standard deviation (SD, given with its standard error) of the ratio MIC MGIT 960/MIC agar dilution, expressed as a number of 2-fold dilutions and denoted by σ. The geometric SD was computed by fitting a log-normal distribution to the ratio MIC MGIT 960/MIC agar dilution as implemented in the R package fitdistrplus (v.1.0-9) (Delignette-Muller *et al.*, 2015). If the data was compatible with $\sigma = 0$, the geometric standard deviation could not be estimated and was defined as "not applicable" (NA). The approach is a generalization of the Bland and Altman method (Martin Bland *et al.*, 1986), taking censoring of the data into account. Strains for which the MGIT 960 MIC and 7H10 agar dilution MIC were both left-censored or both right-censored were excluded since no information on the ratio could be derived. Goodman and Kruskal's gamma was used to quantify the rank correlation between the two methods. No correlation could be calculated if the variance for either method was 0 (NA).

Distributions of wt and mutant MICs were analysed qualitatively based on the results of 7H10 agar dilution. We divided the dataset into two groups: drugs for which the MIC distributions of wt and mutant strains did not overlap, and those for which MIC distributions overlapped.

Sensitivities and specificities of WGS-based resistance profile inference were calculated based on the 7H10 agar dilution results for all drugs, except pyrazinamide– for which the MGIT 960 results were used, based on resistance/susceptibility at the ECOFF defined derived from our data and the presence or absence of a putative resistance-associated mutation.

4.3.4. Defining clinical breakpoints for high/low-level resistance

The therapeutic window of a drug is defined as the maximal serum concentration which is considered safe (Böttger, 2011). Mutations can increase the MIC beyond the therapeutic window and render the drug clinically ineffective. Drugs may have large therapeutic windows beyond the ECOFF. For these, MIC increases caused by mutations may still be within the therapeutic window of a drug: these strains might still be treatable by increasing the drug dose. We analysed the distribution of MICs of mutant strains, and assessed if cut-offs for low-level (within the therapeutic window) and high-level (beyond the therapeutic window) resistance were definable. There were sufficient data available to define distinct low/high-level clinical breakpoint concentrations for isoniazid, rifampicin, streptomycin and amikacin. For mutations conferring resistance to other drugs assayed in this study, no distinct separation into high/low-level resistance was possible due to wide ranges of MICs conferred by the individual mutations or the mutations conferred MICs beyond the therapeutic window.

4.3.5. WGS and single nucleotide variant (SNV) calling

WGS and data analysis was performed as previously described (Ghielmetti *et al.*, 2017) and summarised in the supplementary materials. The performance of WGS-based DST greatly depends on the availability of robust markers of resistance. We therefore focussed on a set of high-confidence resistance-associated genes (Böttger, 2011; Gygli *et al.*, 2017; Walker *et al.*, 2015) (Table 4.2). We additionally assessed the impact of eis promotor mutations on amikacin and capreomycin resistance, as the association of mutations in the eis promotor with resistance to the aforementioned drugs has been reported but is not well established (Kambli *et al.*, 2016; World Health Organization, 2018b).

4.3.6. Ethics

Local institutional review board or ethics committee approval was obtained at all local study sites. Informed consent was obtained where requested per local regulations. This project was approved by the Swiss Ethics Committee on research involving humans (swissethics, Bern, Switzerland).

4.4. Results

4.4.1. Agreement between MGIT 960 and 7H10 agar dilution phenotypic DST

Table 4.3 and Figure 4.1 summarize the agreement between the semi-quantitative and quantitative MIC determination by MGIT 960 and 7H10 agar dilution in terms of classifying strains as resistant or susceptible according to ECOFF (Table 4.1). Agreement was high for all drugs, except ethambutol (see below). For most drugs, the MGIT 960-based MICs were higher than the 7H10 agar dilution-based MICs. MICs obtained using the two methods were within 1-2 two-fold dilution steps of each other. The classifications into resistant or susceptible demonstrated high rank correlations (Table 4.3 and Figure 4.1), except for capreomycin (Figure A1.4) due to few resistant strains included in the study.

Table 4.1.: **Epidemiological cutoffs (ECOFF) used for 7H10 agar dilution and MGIT 960 phenotypic DST.** Values derived from wild-type MIC distributions determined in this study. The values given in parentheses are the critical concentrations recommended by the WHO in 2014 (World Health Organization, 2014a)

Antibiotic	ECOFF agar dilution (mg/L)	ECOFF MGIT 960 (mg/L)
Ethionamide	1 (5)	5
Ethambutol	2 (5)	5
Capreomycin	4	2.5
Streptomycin	0.5 (2)	1
Kanamycin A	2 (5)	2 (2.5)
Amikacin	1 (4)	1
Moxifloxacin	0.25 (0.5)	0.25 (0.5)
Isoniazid	0.125 (0.2)	0.1
Rifampicin	0.5 (1)	1
Rifabutin	0.0625	0.1
Pyrazinamide	NA	100

4.4.2. WGS and *in silico* resistance profile prediction

A total of 176 whole genome sequences with a median coverage of 67.6x (interquartile range (IQR) = 37.48) were obtained. Median mapping percentage and percentage of genome covered were 98.7 % (IQR = 0.94) and 99.4 % (IQR = 0.4), respectively. Genes involved in drug resistance demonstrated high coverages with only 0.8 % of all positions

suffering from coverages below 7x (see supplementary materials). All major *M. tuberculosis* Lineages, except Lineage 7, were represented in the study (L1 = 6, L2 = 36, L3 = 11, L4 = 123, L5 = 1, L6 = 1). The strains showed a range of drug resistance profiles (Figure 4.2). Based on the set of analysed genes (Table 4.2), 25 strains were predicted to be fully susceptible against all assayed drugs, 59 strains were mono-/poly-resistant, 91 strains demonstrated MDR phenotypes and two strains were extensively drug-resistant (XDR: isoniazid, rifampicin, fluoroquinolone and aminoglycoside resistant).

4.4.3. Drug resistance profile prediction by WGS vs. phenotypic DST

After exclusion of known phylogenetic markers not involved in resistance, WGS-based prediction of drug resistance using a defined set of target genes (Table 4.2) was highly congruent with the categorical classification based on the phenotypic DST for most drugs (Table 4.3, Table 4.4, Figure 5). Based on the *in silico* resistance prediction, the MICs of mutant and wt strains frequently followed a Gaussian distribution. However, the same resistance marker may confer different MICs in different strains (Figures A1.1C, A1.2C, A1.3C, A1.8C, A1.9C A1.10C). In some cases, the increase in the MIC conferred by a certain resistance mutation fell within the distribution of the wt MIC (e.g. for *gidB*, *eis* promotor mutations, Figures A1.3C A1.6C).

4.4.4. Distinct wt and mutant MIC distributions

MIC distributions of wt and mutant strains were well separated for rifampicin, rifabutin, isoniazid, kanamycin A, amikacin, capreomycin, streptomycin and pyrazinamide, indicating that the resistance markers used had a high positive predictive power (88.9 % overall positive predictive power of resistance markers). For streptomycin, two strains harboured no mutations in the target genes, yet demonstrated high-level phenotypic resistance (Figure A1.3C).

4.4.5. Overlapping wt and mutant MIC distributions

MIC distributions of wt and mutant strains overlapped for ethambutol, moxifloxacin and ethionamide (Figure 4.3). For ethambutol and ethionamide, overlapping MIC distributions of wt and mutant strains were associated with a large number of polymorphisms

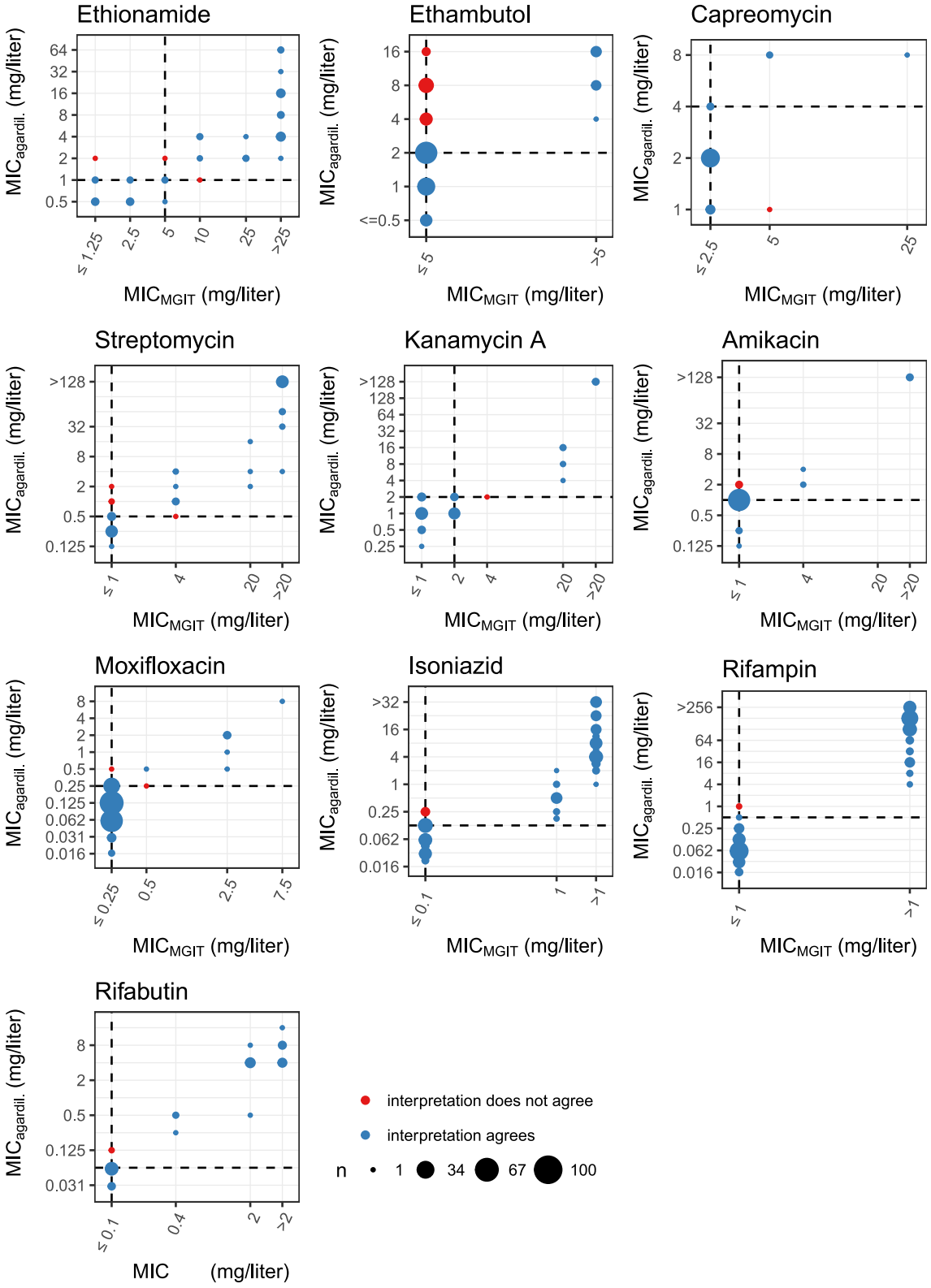

Figure 4.1.: **Method agreement between phenotypic DST performed with MGIT 960 and 7H10 agar dilution represented as Bland-Altman plots for all drugs tested in this study.**

in resistance-conferring genes (ethambutol resistance: 22 polymorphisms in *embB*, ethionamide resistance: 28 in *ethA*, 3 in *inhA*, 6 in *inhA* promoter). Solubility issues with ethionamide led to quantitative differences in MGIT 960 vs. 7H10 agar dilution-based DST (Table 4.3, Figure 5). The overlap in MIC distributions between wt and strains carrying an *embB* mutation was reduced by adjusting the critical concentration for ethambutol resistance from 5 mg/L to 2.5 mg/L (MGIT 960). However, there was variability in the MICs for the same mutation (e.g. MIC EmbB M306I/V in 7H10 agar dilution: 4-16 mg/L – Figure A1.2C). Moxifloxacin resistance was rare (n = 9, MGIT 960, critical concentration 0.25 mg/L) and MIC distributions of mutant strains partially overlapped with those of wt. Sensitivity of the genome-based moxifloxacin resistance prediction 80.0 % (Table 4.1).

Table 4.2.: **List of genes implicated in drug resistance in *M. tuberculosis*.** Genes were screened for polymorphisms by WGS. List adapted from (Gygli *et al.*, 2017)

Drug	**Target gene(s)**
Ethionamide	*ethA*, *inhA*, *inhA* promoter
Ethambutol	*embB*
Capreomycin	*rrs, eis promoter, tlyA*
Streptomycin	*rrs, gidB, rpsL*
Kanamycin A	*rrs, eis promoter*
Amikacin	*rrs, eis promoter*
Moxifloxacin	*gyrA*
Isoniazid	*katG, inhA promoter*
Rifampicin/rifabutin	*rpoB*
Pyrazinamide	*pncA, pncA promoter*

Table 4.3.: **Summary statistics of the method agreement between 7H10 agar dilution- and MGIT 960-based phenotypic DST for all drugs assayed in this study.**

Antibiotic	n	Categorical agreement (%)	SD of $\log_2$ (MIC MGIT 960/MIC agar dilution)	γ
Ethionamide	56	95	1.9 ± 0.3	0.91
Ethambutol	171	73	1.9 ± 0.5	0.94
Capreomycin	56	98	1.5 ± 0.5	0.65
Streptomycin	56	93	1.5 ± 0.3	0.98
Kanamycin A	56	98	1.2 ± 0.2	0.8
Amikacin	174	98	1.4 ± 0.6	1
Moxifloxacin	173	99	1 ± 0.2	1
Isoniazid	173	96	1.2 ± 0.1	1
Rifampicin	174	99	NA	1
Rifabutin	56	96	0.8 ± 0.1	0.98

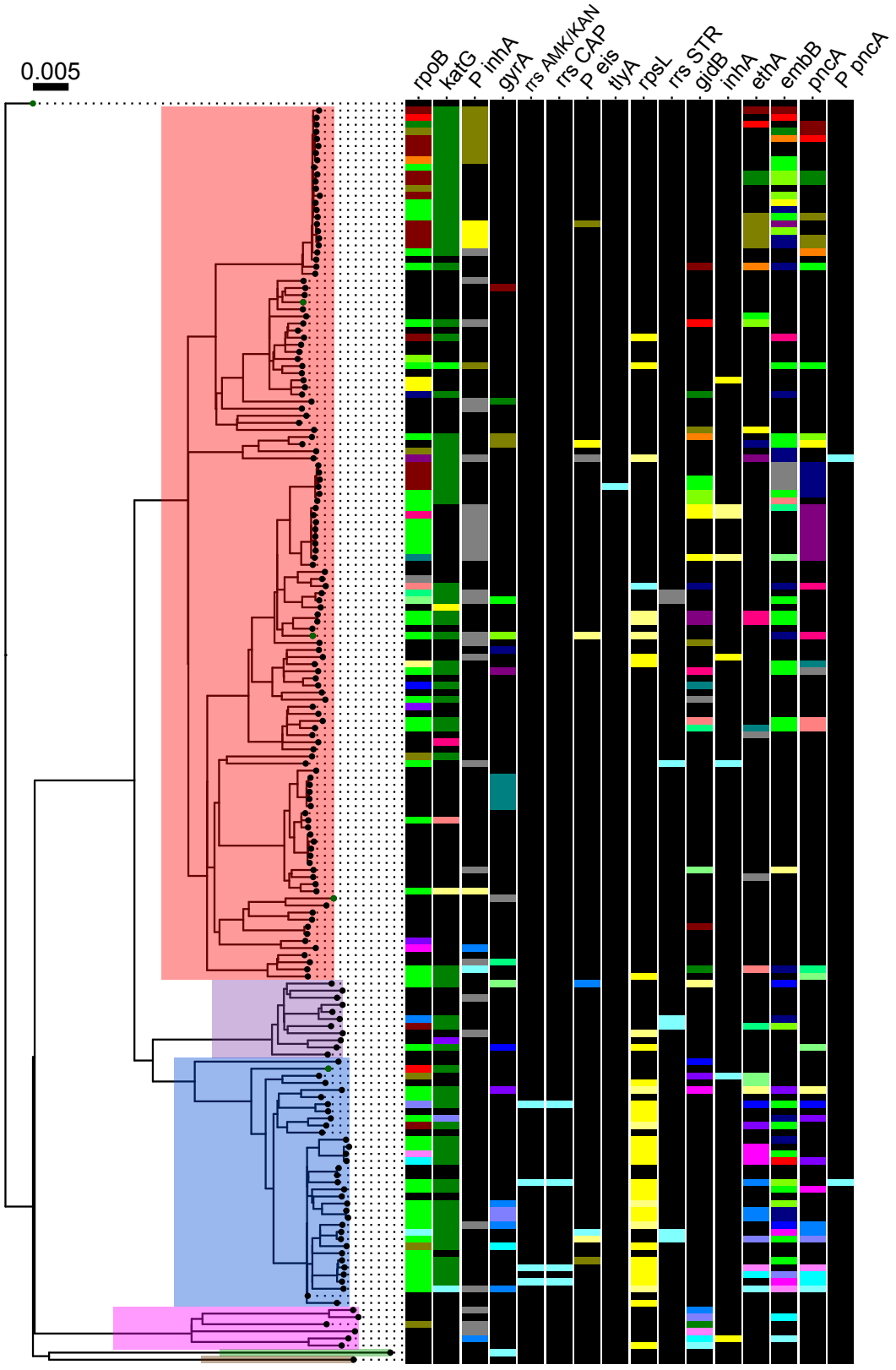

Figure 4.2.: **Phylogeny of 176 *M. tuberculosis* strains based on 20510 variable positions.** Green tip labels depict reference strains. Main Lineages are highlighted as follows: Red L4, purple L3, blue L2, pink L1, green L6, brown L5. Scale bar: substitutions per site. Phylogeny rooted on *M. canettii*. Colored bars indicate mutations per gene and within a distinct column (gene) each colored bar represents a distinct mutation. Black bars indicate wt.

4.4.6. Defining high-/low-level clinical breakpoint concentrations

Isoniazid

Mutations in the promoter of *inhA* conferred low-level resistance < 1 mg/L (7H10 agar dilution), compared to strains harbouring mutations in *katG* or combinations of *inhA* promoter and *katG* mutations which demonstrated MIC levels ranging from > 1 mg/L to > 32 mg/L in 7H10 agar dilution (Figure A1.8.C). Defining clinical breakpoint concentrations (CBC) for low- (≤ 1 mg/L for MGIT 960/7H10 agar dilution) and high-level (> 1 mg/L MGIT 960/7H10 agar dilution) isoniazid resistance is warranted.

Rifampicin/Rifabutin

Most mutations in *rpoB* increased the MIC for rifamycins beyond the therapeutic window (peak serum concentration 10 mg/L (Böttger, 2011; Sekaggya-Wiltshire *et al.*, 2018)). However, some rare *rpoB* mutations (e.g. RpoB L452P, H445L – Figure A.1.9.C) demonstrated MICs within the therapeutic window. Defining low- and high-level CBC may thus be justified. For rifampicin, CBC were ≤ 4/2 mg/L for MGIT 960/7H10 agar dilution and > 4/2 mg/L for MGIT 960/7H10 agar dilution, respectively. For rifabutin, our data suggests CBC for low- and high-level resistance of ≤ 0.4/0.25 or 0.5 mg/L for MGIT 960/7H10 agar dilution and > 0.4/0.25 or 0.5 mg/L for MGIT 960/7H10 agar dilution, respectively. Mutations in *rpoB* conferring resistance to rifampicin and rifabutin showed highly correlated increases (Figure 4.4) of MICs beyond the therapeutic window for most *rpoB* mutations 3 and Figure A.1.9.C & A.1.10.C), indicating that both drugs are rendered clinically ineffective by the mutations identified in the dataset (Berrada *et al.*, 2016) and cannot substitute each other.

Amikacin

Few strains had mutations in the regions of rrs relevant for amikacin resistance or the eis promoter (n=12). Mutations in *rrs* were associated with high-level (> 128 mg/L in 7H10 agar dilution) resistance. With regards to the *eis* promotor, only the C-14T mutation increased the MIC and leads tos low-level (2-4 mg/L in 7H10 agar dilution) amikacin resistance. Given the peak serum concentrations of amikacin (20-40 mg/L (Böttger, 2011)), a CBC for low- (≤ 4 mg/L for MGIT 960/7H10 agar dilution) and high-level (4 mg/L for MGIT 960/7H10 agar dilution) amikacin resistance may be warranted.

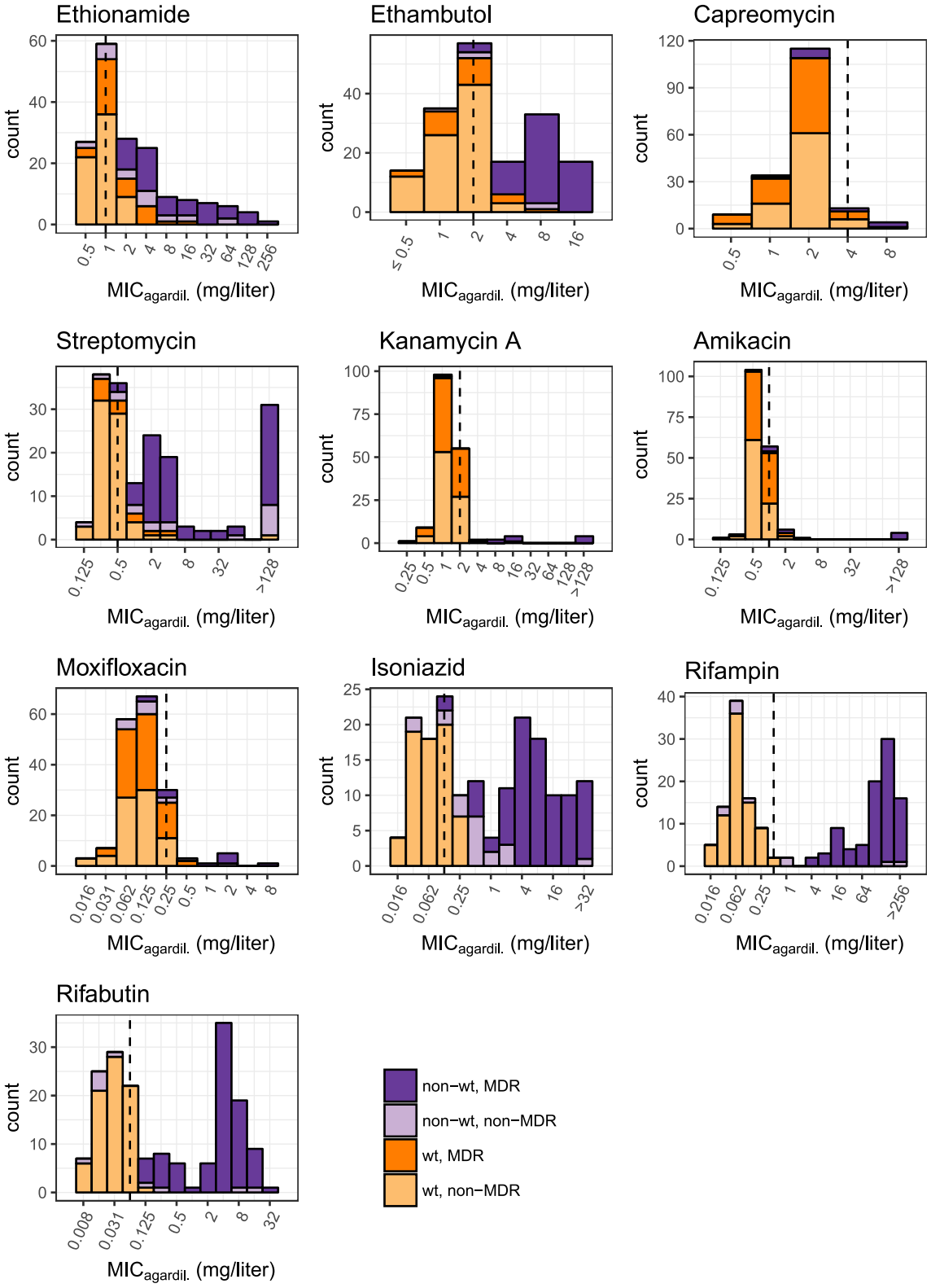

Figure 4.3.: **Histograms of MICs of (7H10 agar dilution) for all drugs assayed in this study.** Dashed bar indicates the epidemiological cutoff.

Streptomycin

Certain mutations lead to MICs well beyond the peak serum concentrations (Böttger, 2011) of streptomycin (e.g. RpsL K43R, MIC 7H10 agar dilution > 128 mg/L, Figure A.1.3.C). On the other hand, *gidB* mutations increase the MIC only moderately (MIC 7H10 agar dilution 1-4 mg/L, Figure A.1.3.C). Mutational combinations in *gidB, rrs, rpsL* were common and produced a range of different MICs. However, there were mutations that systematically lead to MICs beyond the therapeutic window, e.g. RpsL K43R. Defining low- level (MGIT 960 ≤ 4 mg/L, 7H10 agar dilution ≤ 4-8 mg/L) and high-level CBC for streptomycin resistance (MGIT 960 > 4 mg/L, 7H10 agar dilution > 4-8 mg/L) is warranted.

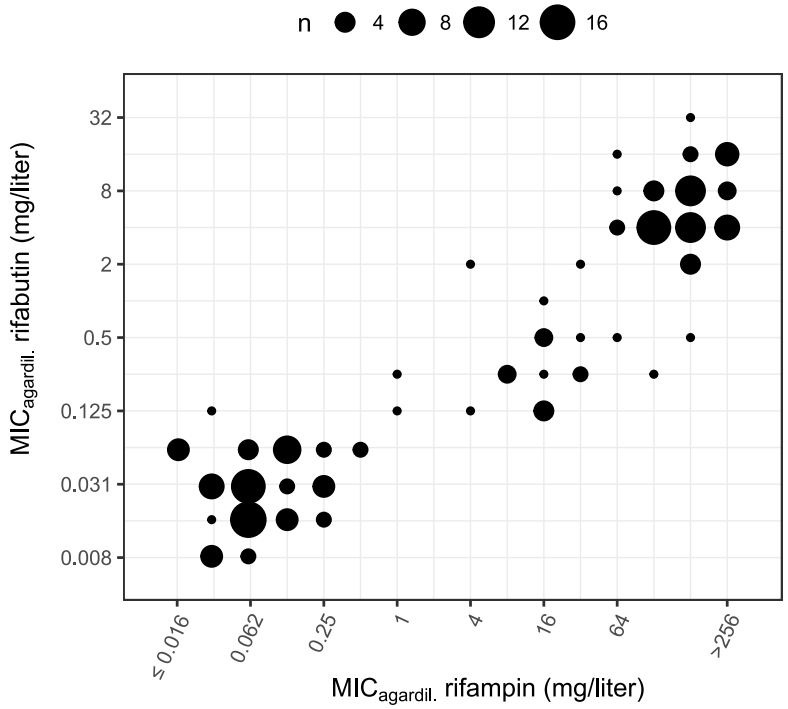

Figure 4.4.: **Correlation between MICs for 7H10 agar dilution rifampicin and rifabutin.**

4.5. Discussion

We compared quantitative phenotypic DST with *in silico* or genomic resistance profile prediction inferred from WGS using 176 clinical *M. tuberculosis* isolates. We observed a bias towards Lineage 4 and 2 strains in our sample set. The initial strain set used to

establish the methodology was collated with a specific aim to include drug-resistant strains and given the frequent association of Lineage 2 and 4 with drug-resistance (Casali *et al.*, 2014; Cohen *et al.*, 2015), the observed bias is not surprising. Furthermore, Lineage 2 and 4 strains are also frequently isolated at the collection sites of the strain set used to apply the methodology (Ivory Coast, Peru and Democratic Republic of the Congo). Similarly, resistance against a number of drugs (amikacin, capreomycin, kanamycin, moxifloxacin) was rare, reflecting the scarcity of pre-XDR/XDR phenotypes in Switzerland (initial strain set) and at the sites of prospective sampling.

The results of MGIT 960 and 7H10 agar dilution-based phenotypic DST methods were highly correlated and suitable to separate susceptible from resistant variants. Based on phenotypic DST results and WGS, we were able to define CBC for high- and low-level resistance for isoniazid, rifampicin, streptomycin and amikacin. Defining such breakpoints is important for preserving efficacious drugs for treatment of resistant *M. tuberculosis* variants. Our data suggest that the current WHO-defined critical concentration for phenotypic DST of ethambutol by MGIT 960 (5 mg/L) is too high and may misclassify strains as susceptible when compared to the 7H10 agar dilution-based classification. Given the low peak serum concentrations for ethambutol, this may lead to mistreatment due to presumed ethambutol susceptibility. After adjusting the critical concentration to 2.5 mg/L for MGIT 960, we observed a strong improvement of the categorical agreement between MGIT 960- and 7H10 agar dilution-based classification.

Table 4.4.: **Sensitivity and specificity of the genome-based drug resistance profile prediction.** The 7H10 agar dilution-based categorical classification was used as the gold standard for all drugs except pyrazinamide, for which the MGIT 960 categorical classification was used.

Drug	**Sensitivity (%)**	**Specificity (%)**
Ethionamide	75.0	92.9
Ethambutol	89.6	94.2
Capreomycin	75.0	94
Streptomycin	68.0	92.1
Kanamycin A	83.3	98.8
Amikacin	63.6	96.9
Moxifloxacin	80.0	90.2
Isoniazid	93.6	96.8
Rifampicin	100	94.0
Rifabutin	98.9	94.0
Pyrazinamide	80.8	88.9

The mutations identified by WGS had a high predictive power to classify strains as resistant. However, the predictive power depends on a number of factors. For instance, the increase in MIC conferred by an identical resistance mutation can vary greatly in different strains (e.g. EmbB M306I/V, RpsL K88R) (Ruesen *et al.*, 2018). Such variation is clinically relevant if there is a significant overlap between the MICs of mutant and wt strains (Schön *et al.*, 2017), as was the case for ethionamide, ethambutol and streptomycin (e.g. *gidB*) resistance mutations. Furthermore, it is difficult to classify strains as resistant or susceptible if the MIC increase lies within the therapeutic window of a drug. The overlap between MICs of mutant and wt strains is confounded by the fact that we only screened for mutations in genes which had previously been associated with drug resistance. We might thus have missed possible resistance-conferring mutations in other genes. Additionally, WGS will always produce distributions of coverages which in term will inevitably lead to certain regions in the genome suffering from low coverage, preventing the detection of mutations. However, in cases where we observed elevated MICs without any mutations detected the target genes, coverage issues could not explain the absence of any mutations.Furthermore, the strain genetic background (Fenner *et al.*, 2012), non-mutational mechanisms (e.g. modulation of gene expression) (Freihofer *et al.*, 2016), as well as drug efflux mechanisms (Silva *et al.*, 2016) may contribute to the variability in increase of the MIC conferred by resistance mutations.

The predictive power of mutations in target genes also depends on removing phylogenetic markers not involved in resistance. Separating phylogenetic from resistance-associated markers works well for essential (highly conserved) genes such as *rpoB, rpsL, rrs* but is problematic in non-essential genes involved in the conversion of prodrugs into their active forms like *pncA* (pyrazinamide), *ethA* (ethionamide) or in genes that generally exhibit higher numbers of polymorphisms e.g. *embB*. Of note, the *embABC* operon is highly polymorphic, harbouring more polymorphisms than expected by chance (mutations in *embABC* operon = 81, expected = 44.8, p = 9.174e-07, binomial test). Mutations conferring ethambutol resistance (Safi *et al.*, 2013) will therefore inevitably evolve in the presence of phylogenetic SNVs and may interact epistatically to produce the variability in MICs we observed for wt strains and for the most common ethambutol resistance markers EmbB M306I/V. The *embABC* operon is involved in the biosynthesis of decaprenylphosphoryl-β-d-arabinose, which is an integral component of the mycobacterial cell wall. The cell envelope interacts with the host immune system and the high levels of diversity of these genes might be the product of diversifying selection due to host immune pressure. The influence of polymorphisms in the *embABC* operon on MICs in general is supported by the observation that sub-inhibitory concentrations of ethambutol lower

the MICs for isoniazid, rifampicin and streptomycin (Jagannath *et al.*, 1995). Even small changes in activity of the decaprenylphosphoryl-β-d-arabinose biosynthetic and utilisation pathway might thus alter cell wall permeability and influence MICs of several drugs.

Similarly, in the case of streptomycin resistance, the RpsL substitution K88R exhibited a range in MICs from low to high-level resistance making it difficult to judge the susceptibility of a strain harbouring this mutation based on the genotype. Streptomycin was the first effective antituberculous drug discovered (Schatz *et al.*, 1944) and has been used for decades. The long-term use has produced complex resistance profiles with multiple streptomycin resistance mutations (e.g. in *gidB*, *rpsL*, *rrs*) occurring concomitantly, producing wide ranges of MICs. Furthermore, many streptomycin resistant strains displayed MDR/XDR phenotypes. Streptomycin resistance mutations are frequently found in backgrounds which have mutations in genes affecting the information pathway (DNA → RNA → proteins) – e.g. *gyrA* (DNA gyrase), *rpoB* (DNA-dependent RNA polymerase), *rrs* (ribosomal RNA). The simultaneous presence of multiple resistance mutations may alter the adaptive landscape (Borrell *et al.*, 2013; Moura de Sousa *et al.*, 2017). In addition, non-mutational processes (e.g. alteration of gene expression) may compensate for fitness costs of drug resistance and at the same time alter the MIC for the drug (Freihofer *et al.*, 2016). This has not been demonstrated for streptomycin resistance in *M. tuberculosis*, but it seems possible that compensation of fitness costs in MDR phenotypes might alter the MIC for streptomycin (Moura de Sousa *et al.*, 2017), considering that streptomycin is not part of the current standard treatment regimen and selection for high-level streptomycin resistance is relaxed.

With 63.6 % – 80.8 %, sensitivities were low for a number of drugs (i.e. for amikacin, moxifloxacin, pyrazinamide) (table 4.4) (World Health Organization, 2018c), but were comparable to other study not employing a database of pre-defined resistance mutations (Farhat *et al.*, 2016; Shea *et al.*, 2017; Walker *et al.*, 2015). The observed low sensitivities for some drugs were either due to few resistant strains included in the dataset, the presence of additional resistance mutations in genes not assessed or due to unknown resistance mechanisms and not due to low coverages prohibiting the detection of mutations. The use of a curated SNV-database containing high-confidence drug-resistance mutations would improve sensitivity for some drugs where additional targets, less well associated with resistance, are known (Miotto *et al.*, 2017; World Health Organization, 2018c). However, reliance on a predefined resistance mutation database comes at the cost of reduced specificity. After known phylogenetic mutations have been removed, it is important to treat any mutation in known target genes as potentially involved in drug resistance. In

cases where previously unknown mutations (i.e. neither known to confer resistance, nor a known phylogenetic SNV) in resistance-related genes are detected, targeted DST is necessary to confirm or reject the drug-resistance conferring nature of a novel mutation to achieve high sensitivities and specificities for whole genome sequencing-based DST. Generating high-quality quantitative DST data using diverse *M. tuberculosis* strains is important to accurately define the ECOFF and subsequently guide treatment decisions. The two quantitative DST methods employed are difficult to standardize across laboratories, technically demanding and at best challenging to scale up. Microtiter plate-based quantitative DST methods (Lee *et al.*, 2014; Rancoita *et al.*, 2018) have the potential to aid in the generation of more high quality DST data due to their standardized formulation and relative ease of application compared to established methods.

In conclusion, we demonstrate that MGIT 960 and 7H10 agar dilution-based phenotypic DST provide highly congruent classifications of strains into resistant or susceptible. WGS has high predictive power to infer resistance profiles without the need for time-consuming phenotypic methods. Limitations due to overlapping distributions of wt and mutant MICs, varying MICs for the same resistance mutations in different strains, presence of phylogenetic markers in resistance-associated genes and rare resistance markers with low frequencies will likely be resolved by on-going large-scale projects (e.g. ReSeqTB and others (Starks *et al.*, 2015; The CRyPTIC Consortium and the 100000 Genomes Project, 2018)) combining phenotypic DST with WGS of thousands of *M. tuberculosis* isolates. Our findings, together with those of on-going studies will pave the way for the replacement of phenotypic DST with drug resistance profile prediction based on WGS in the coming years.

4.6. Acknowledgements

We would like to thank Alexandra Mushegian for critically reading the manuscript and improving the writing. Whole genome analysis was performed at sciCORE scientific computing core facility at University of Basel (http://scicore.unibas.ch/).

4.7. Funding

This work was supported by the European Research Council [grant number 309540-EVODRTB], the Swiss National Science Foundation [IZRJZ3_164171, 310030-166687, IZLSZ3_170834, CRSII5_177163, 31003A_153349 and 320030_153442/1] and

SystemsX.ch. The International epidemiology Databases to Evaluate AIDS (IeDEA) is supported by the National Institutes of Allergy and Infectious Diseases (NIAID) under award numbers U01 AI096299, U01 AI069919, U01 AI069924, U01 AI069911, U01 AI069907, U01 AI096186, and U01 AI069923. The content of this publication is solely the responsibility of the authors and does not necessarily represent the official views of any of the governments or institutions mentioned above.

4.8. Conflict of interest

Peter M. Keller reports travel grants by Copan Italia SpA outside of the submitted work. Erik C. Böttger is a consultant for AID Diagnostics.

5. Prisons as ecological drivers of fitness-compensated and multidrug-resistant *Mycobacterium tuberculosis*

Sebastian M. Gygli[1,2], Levan Jugheli[1,2,3], Natia Adamia[3], Chloé Loiseau[1,2], Andrej Trauner[1,2], Miriam Reinhard[1,2], Amanda Ross[1,2], Sonia Borrell[1,2], Rusudan Aspindzelashvili[3], Nino Maghradze[1,2,3], Klaus Reither[1,2], Zaza Avaliani[3], Sebastien Gagneux[1,2]

[1]Swiss Tropical and Public Health Institute, Basel, Switzerland
[2]University of Basel, Basel, Switzerland
[3]National Center for Tuberculosis and Lung Diseases (NCTLD), Tbilisi, Georgia

5.1. Abstract

Multidrug-resistant tuberculosis accounts for one third of the annual deaths due to antimicrobial resistance (O'Neill, 2014). Drug-resistance conferring mutations frequently cause fitness costs in bacteria (Andersson *et al.*, 1999; Andersson *et al.*, 2010; Gagneux *et al.*, 2006a; Gygli *et al.*, 2017). It is therefore widely believed that multidrug-resistant tuberculosis is mainly due to *de novo* evolution of resistance linked to patient non-adherence rather than transmission (Frieden *et al.*, 2007). Experimental work indicates that drug resistance-related fitness costs may be mitigated by compensatory mutations (Comas *et al.*, 2012; Maisnier-Patin *et al.*, 2002; Qi *et al.*, 2014; Reynolds, 2000; Song *et al.*, 2014). Yet, the clinical relevance of compensatory evolution and its ecological determinants remain poorly understood. Here, we show that in the Republic of Georgia, at least 43 % of multidrug-resistant tuberculosis was due to patient-to-patient transmission. Transmission was independently associated with compensatory mutations and a patient history of incarceration. Compensatory mutations were over-represented in prison-derived strains that spilled over into the general population. We conclude that prisons fuel the epidemic of multidrug-resistant tuberculosis in the Republic of Georgia by acting as ecological drivers of highly transmissible strains.

5.2. Main Text

Growing antimicrobial resistance is a threat to global public health and the economy (O'Neill, 2014). In 2016, there were an estimated 560,000 new cases of human tuberculosis (TB) caused by multidrug-resistant (MDR) *Mycobacterium tuberculosis* (*Mtb*) strains resistant to the two first-line antibiotics isoniazid and rifampicin. However, the number of fatalities is small compared to the annual total of 1.6 million deaths due to TB in general (O'Neill, 2014; World Health Organization, 2018a). Globally, only 4-6 % of the 10.0 million annual new TB cases are caused by MDR *Mtb* variants, and this proportion has remained stable despite the fact that TB has been treated with antibiotics for decades (Dye, 2009). Based on these observations, it has been suggested that MDR-TB might be generally less transmissible due to fitness costs of MDR (Dye *et al.*, 2001), and as a consequence, MDR-TB was predicted to remain a localized problem (Dye *et al.*, 2002). Indeed, several geographical hotspots of MDR-TB exist, with the countries of the former Soviet Union being heavily affected, for reasons not well understood. For instance, in the Republic of Georgia, 11 % of all new TB cases and 30 % of retreatment cases in 2017 were caused by MDR strains (World Health Organization, 2018a).

In *Mtb*, drug resistance is mainly conferred by chromosomal mutations in the genes encoding the drug target (Gygli *et al.*, 2017). *In vitro* data demonstrated a fitness deficit for rifampicin-resistant *Mtb* (Gagneux *et al.*, 2006c). In contrast, analysis of paired clinical isolates from patients who acquired rifampicin resistance during treatment revealed that some of these strains did not carry any detectable fitness deficit *in vitro*. It was hypothesized that these clinical strains acquired secondary, fitness-deficit compensating mutations. Genome analyses of experimentally evolved *Mtb* laboratory strains together with a collection of rifampicin-resistant clinical strains revealed the presence of compensatory mutations in the RNA polymerase (Comas *et al.*, 2012). Work conducted in several bacterial species (Brandis *et al.*, 2012; Reynolds, 2000; Song *et al.*, 2014) showed that secondary mutations in the RNA polymerase restored the transcriptional activity of the enzyme. However, whether these compensatory mutations influence the transmission fitness of *Mtb* in human populations remains to be established. While several studies assessed the impact of compensatory mutations on the transmissibility of MDR-TB, findings have been inconsistent (Casali *et al.*, 2014; Liu *et al.*, 2018; Merker *et al.*, 2018; Vos *et al.*, 2013). Moreover, these previous studies relied on convenience sampling and did not control for confounding factors. In this study, we tested for associations between various bacterial and corresponding patient data with the transmission of MDR-TB isolates by whole-genome analysis of a nation-wide collection of 659 *Mtb* strains

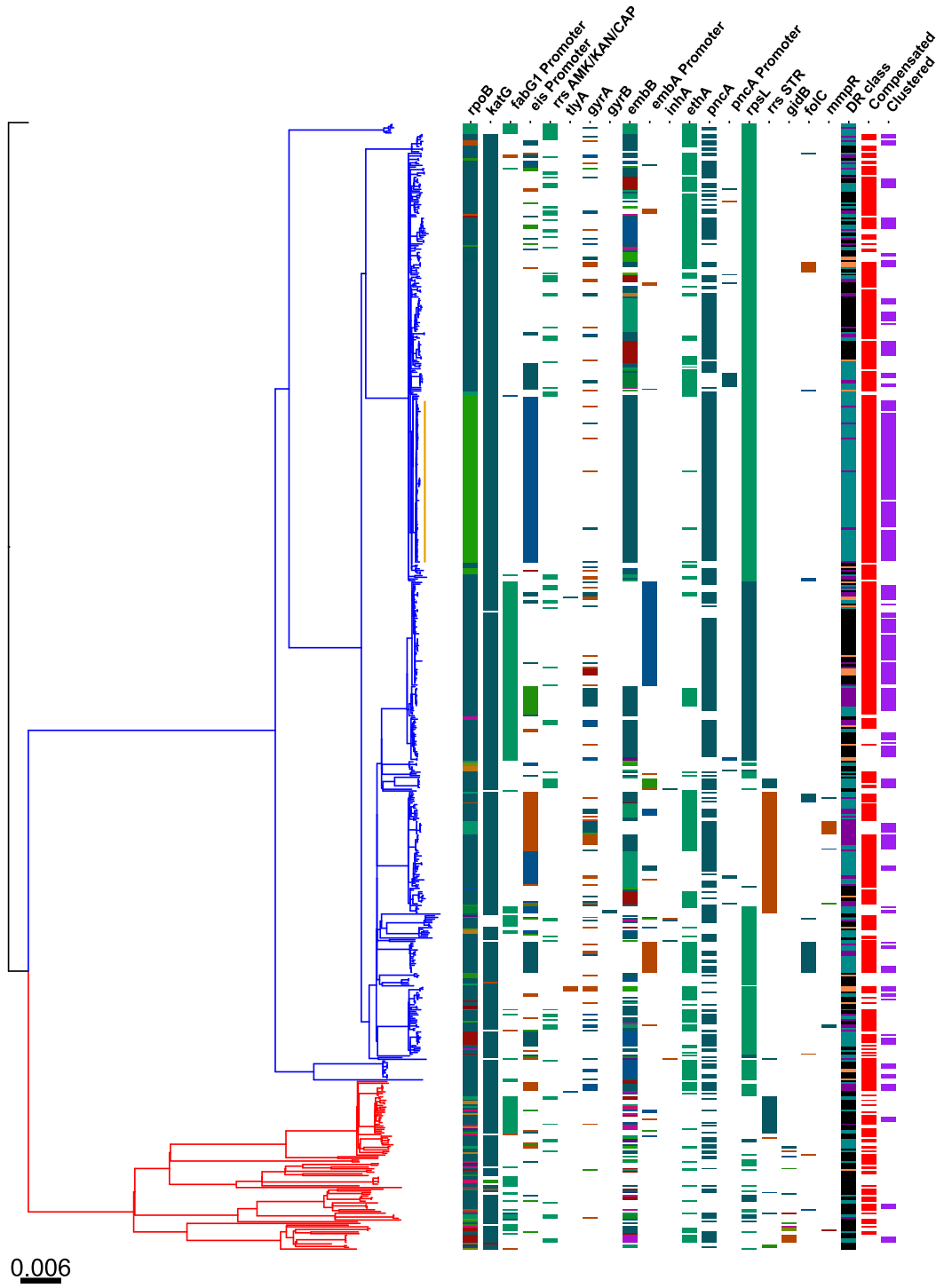

Figure 5.1.: **Phylogeny of 659 MDR *M.tuberculosis* strains together with the drug resistance mutations per strain.** Blue branches correspond to Lineage 2 (Beijing sublingeage), red to Lineage 4, the black bar indicates the root (*Mycobacterium canettii*). Scale bar indicates substitutions/site. Each unique mutation in the displayed target genes is coded with a different color, except for *pncA* and *ethA* due to the large number of different mutations. The orange bracket indicates the largest detected transmission cluster with n = 84 clustered strains.

from the Republic of Georgia isolated from 2011 to 2013. The dataset represents 53 % of all culture-confirmed MDR-TB cases isolated in the Republic of Georgia in this timeframe (Figures A.2.1-A2.3).

We first assessed the drug resistance profiles of the sequenced *Mtb* strains and found that the epidemic of drug-resistant TB in Georgia is driven by highly resistant strains (Figure 5.1). Out of the 659 strains analyzed, 304 strains (46 %) were pXDR (pXDR: MDR + resistance to either aminoglycosides or fluoroquinolones), 257 (39 %) were MDR, and 98 (15 %) extensively drug resistant (XDR: MDR + fluoroquinolone and aminoglycoside-resistant). The pXDR and XDR resistance profiles are strongly associated with treatment failure (Ahmad *et al.*, 2018; Lomtadze *et al.*, 2009). Apart from the mutations defining the drug resistance class, all strains carried a median of 3 additional resistance mutations (25^{th} percentile = 3, 75^{th} percentile = 4). Control measures will be hampered by the predominance of pXDR phenotypes and the high prevalence of additional resistance mutations; for example, 92 % of the strains harbored streptomycin resistance-conferring mutations (Figure 5.1). Treatment regimens including the novel drugs bedaquiline, pretomanid, delamanid and linezolid will be necessary to combat the epidemic of MDR-TB in Georgia (WHO, 2014). On that note, we identified 16 strains with mutations in the promoter region or coding sequence of the transcriptional repressor Rv0678/MmpR (Figure 5.1). Mutations in the gene Rv0678/*mmpR* are implicated in bedaquiline/clofazimine cross-resistance (Ismail *et al.*, 2018) and bedaquiline has been used in Georgia on a compassionate basis since 2011 (World Health Organization, 2014b).

To identify putative compensatory mutations we screened RNA polymerase genes *rpoABC* for the presence of nonsynonymous substitutions. After filtering for phylogenetic markers (see Materials and Methods), we identified a total of 99 distinct substitutions (RpoA: N = 11, RpoB: N = 53, RpoC: N = 53, Figure 5.2). Strains belonging to Lineage 2 were more likely to harbor compensatory mutations compared to Lineage 4 ($\chi^2 = 29.16$, $p < 0.001$). Most compensatory mutations evolved only once in the dataset and were only shared among a limited number of strains (Figure 5.2). However, several mutations were shared by a large number of strains and independently evolved multiple times (Figure 5.2), indicating a strong selective benefit when carrying these mutations.

We next used the whole genome sequences to identify transmission clusters based on the genetic distance between two given strains to determine whether MDR *Mtb* strains evolved *de novo* in patients or stem from transmission. We identified 42 transmission clusters with a median size of 3.5 strains (25^{th} percentile = 3, 75^{th} percentile = 5), resulting in 43 % (N = 281) of strains being clustered (Figure 5.1). The high proportion of strains in

clusters indicates frequent transmission of MDR-TB. We detected a cluster containing 84 pXDR strains which is one of the largest transmission clusters of pXDR strains reported to date (Figure 5.1). To test the hypothesis that compensatory mutations contribute to the transmission fitness of MDR *Mtb* in Georgia, we performed multivariable logistic regression on a subset of 544 strains for which complete epidemiological records were available (Fig. A.2.3). Supporting our hypothesis, we found that clustering was independently associated with compensatory mutations in the RNA polymerase (OR_{adj} = 2.27, CI_{95} = 1.38–3.79, p = 0.002; Table 1). Other independent predictors of MDR-TB transmission included a patient history of incarceration (OR_{adj} = 6.55, CI_{95} = 3.98–11.05, p < 0.001; Table 1), in line with the known role of prisons in the epidemic of MDR *Mtb* in the former Soviet Union (Aerts *et al.*, 2000; Kenyon, 2009; Stuckler *et al.*, 2008). Older age (OR_{adj} = 0.99, CI_{95} = 0.97–1.00, p = 0.04; Table 1) was negatively associated with clustering, consistent with the epidemiology of TB in middle-/high-burden countries (Yates *et al.*, 2016). In concordance with previous reports (Merker *et al.*, 2018; Niemann *et al.*, 2010), the Lineage 2/Beijing family of *Mtb* was associated with transmission, supporting the notion that Lineage 2 strains may suffer from smaller drug-resistance related fitness costs (Gagneux *et al.*, 2006c).

Table 5.1.: **Risk factors associated with clustering.** ‡Odds ratios were estimated by multivariable logistic regression adjusting for the presence of compensatory mutations, incarceration, *Mtb* lineage, patient age, patient sex, TB diagnosis in the past, isolation year and the total number of drug resistance mutations.

	Clustering			
	All isolates N= 544		Excluding largest cluster N = 457	
	OR_{adj}‡ (CI_{95})	P-value	OR_{adj}‡ CI_{95}	P-value
Compensatory mut. in *rpoABC*	2.27 (1.38, 3.79)	0.002	1.66 (1.01, 2.76)	0.05
Incarceration	6.55 (3.98, 11.05)	<0.001	4.28 (2.47, 7.53)	<0.001
Lineage 2 strain	3.82 (1.94, 8.08)	<0.001	2.87 (1.47, 6.02)	0.004
Age	0.99 (0.97, 1.00)	0.04	0.99 (0.97, 1.00)	0.09
Sex Male	0.90 (0.56, 1.46)	0.67	0.86 (0.53, 1.41)	0.54
TB diagnosis in the past	0.99 (0.66, 1.49)	0.96	0.97 (0.63, 1.49)	0.89
Isolation year 201	1.29 (0.79, 2.11)	0.54	1.3 (0.77, 2.2)	0.6
Isolation year 2013	1.06 (0.63, 1.78)	0.54	1.1 (0.64, 1.91)	0.6
Total number of DR mut.	1.22 (1.06, 1.41)	0.006	1.27 (1.09, 1.43)	0.002

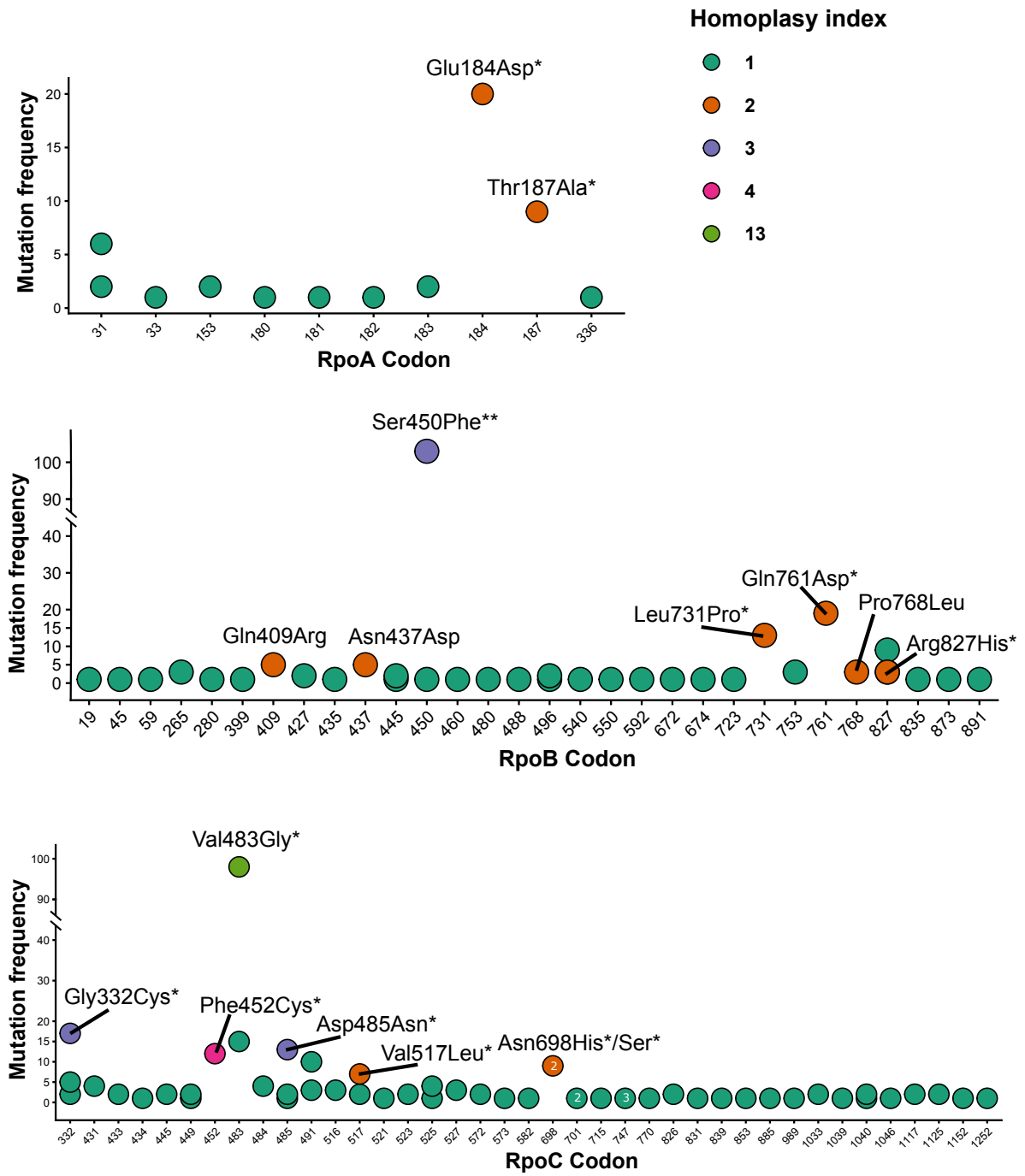

Figure 5.2.: **Putative compensatory mutations.** Mutations in *rpoABC* remaining after filtering for phylogenetic markers. The homoplasic index indicates the number of independent evolution events of the mutation in question. Only mutations with a homoplasy index > 1 are annotated. The frequency of the mutation indicates the number of strains harbouring the respective mutation. Numbers within the dots signify that there were > 1 distinct substitutions affecting the same codon with identical numbers of strains harbouring the mutations. *Mutation has been reported previously (Comas *et al.*, 2012; Merker *et al.*, 2018; Casali *et al.*, 2014). **Strain carrying multiple mutations affecting the same codon.

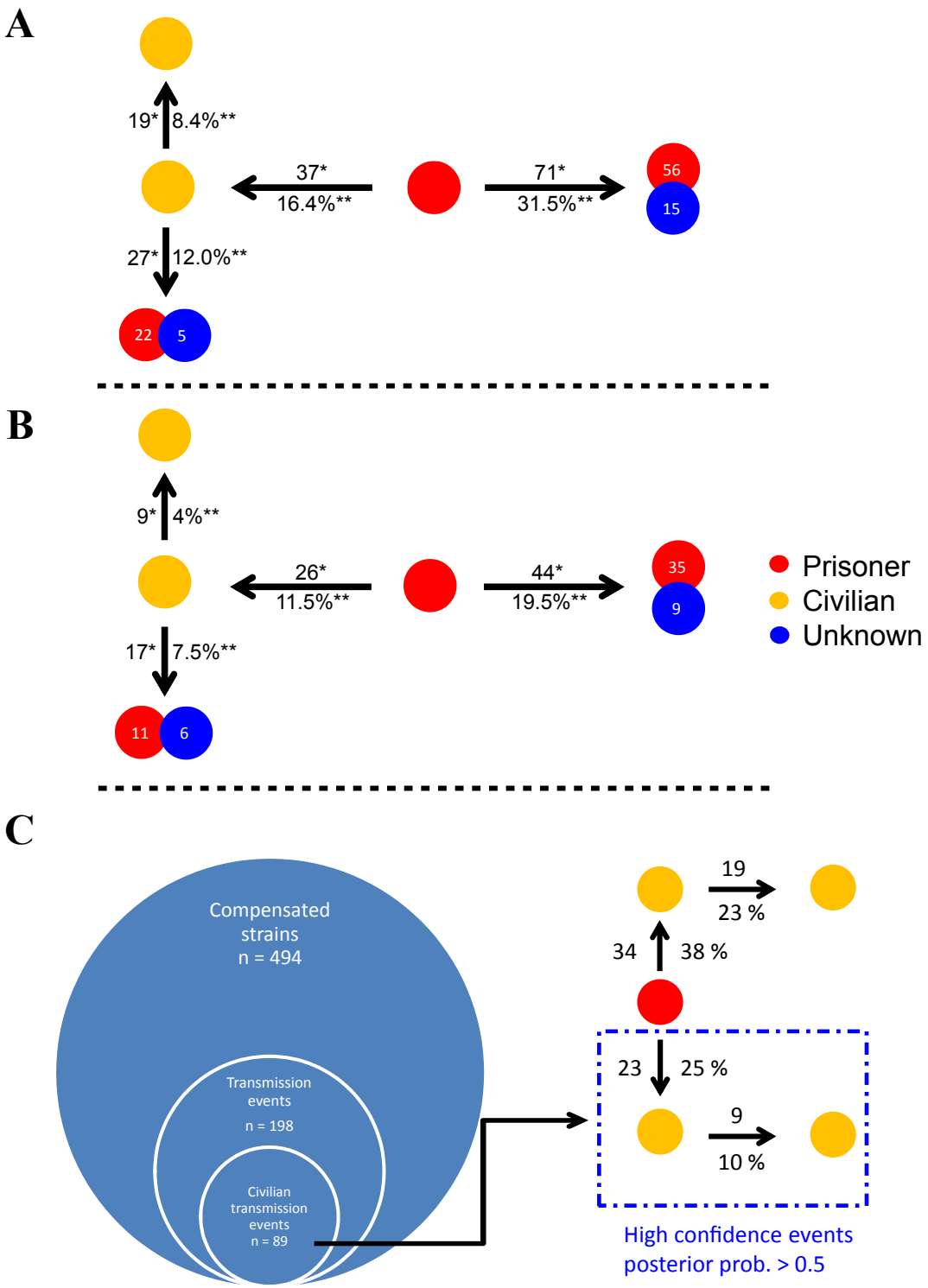

Figure 5.3.: **Summary of the transmission network analysis given in numbers of transmission events.** Arrows indicate the directionality of the transmission event. * Number of transmission events. ** Proportion of the total transmission events (n = 225). Panel A shows all transmission events, panel B only shows high confidence transmission events with posterior probabilities > 0.5. Panel C summarises the sources of compensated strains among civilians. Of the total of N = 494 compensated strains, N = 198 were found in transmission clusters, of which N = 89 were isolated from civilians. At least 35 % (n = 32, posterior probability > 0.5) of the transmitted, compensated strains isolated from civilians have their origin in prisons.

Given the number of transmission events within prisons, we hypothesized that prisons serve as a source for MDR-TB among civilians (Mabud *et al.*, 2019; Warren *et al.*, 2018). To test this hypothesis, we first used the whole genome sequences, isolation dates, and information on the time course of *Mtb* infections to reconstruct transmission networks and infer infection chains including the directionality of transmission. Out of the total 225 transmission events, an estimate of 108 (median posterior probability 0.58, 25^{th} percentile = 0.40, 75^{th} percentile = 0.84), or 48 % were directly linked to transmission from prisoners, and an estimated 37 members of the general public were directly infected by prisoners (Figure 5.3 A). When focusing only on high confidence transmission events (posterior probability > 0.5), an estimated 70 cases (31 % of all transmission events) were associated with prisons. We further quantified indirect spillover from prisons to the general public by determining how many secondary cases were generated by members of general public originally infected by prisoners. An estimate of 46 (N = 26 with posterior probability > 0.5, Figure 5.3 B) secondary cases were generated from members of the general public originally infected by prisoners, of which 19 (N = 9 with posterior probability > 0.5, Figure 5.3 B) were members of the general public. Overall, an estimated 154 cases or 68 % of all transmission events were directly or indirectly linked to incarceration. When focusing only on high confidence transmission events, 96 or 43 % were linked to within or out-of-prison transmission.

Given the strong association between transmission and incarceration, we next hypothesized that repeated transmission of highly drug-resistant strains among prison inmates allows for compensatory evolution. Frequent transmission events would then give rise to highly transmissible strains capable of spilling over into the civilian population. To test the hypothesis, we analyzed the prevalence of compensatory mutations among strains circulating in the general public with or without a connection to prisons. In support of our hypothesis, we found that compensatory mutations were more prevalent in the prison-derived (97 %, 53/56) compared to non-prison derived strains (69 %, 33/48) (p = 0.001, Fisher's Exact Test). Furthermore, compensated strains in transmission clusters isolated from civilians frequently originate from incarcerated patients (N = 32 or 35 %, posterior probability > 0.5, Figure 5.3 C). Prisons therefore are not only sources of MDR-TB for the general public (Mabud *et al.*, 2019; Warren *et al.*, 2018), but they also serve as ecological drivers for compensated, highly transmissible strains.

Our study shows that a large proportion of the MDR-TB in Georgia is due to on-going transmission of highly drug-resistant *Mtb* strains and not solely to *de novo* evolution within patients. Moreover, our findings complement previous experimental findings by

confirming that compensatory mutations in the RNA polymerase of *Mtb* contribute to the transmission fitness of MDR strains in a human population. The strong association between transmission of MDR-TB and incarceration is consistent with previous studies (Aerts *et al.*, 2000; Kenyon, 2009; Stuckler *et al.*, 2008), and further highlights the role of prisons in the epidemic of MDR-TB in the former Soviet Republics. Finally, our new findings revealed a link between incarceration and compensatory evolution and suggest that forcing people into a crowded environment, highly conducive of TB transmission, may influence the evolutionary trajectories of highly drug-resistant bacteria.

5.3. Data availability

The raw sequences were deposited at the National Center for Biotechnology Information under the BioProject ID PRJNA503801. All other data is provided in the supplementary materials.

5.4. Acknowledgements

We thank Alexandra Mushegian for critically reading the manuscript and improving the writing. Calculations were performed at sciCORE (http://scicore.unibas.ch/) scientific computing core facility at University of Basel.

5.5. Funding

This work was supported by the Swiss National Science Foundation (grants 310030_166687, IZRJZ3_164171, IZLSZ3_170834 and CRSII5_177163), the European Research Council (309540-EVODRTB) and SystemsX.ch.

5.6. Author contributions

SG and LJ conceived the idea, SG and ZA supervised the project, SMG performed data curation, data analysis and wrote the first draft, LJ supervised data acquisition and curation, NA carried out data acquisition, CL carried out data analysis, AT carried out data analysis, MR carried out data acquisition, AR carried out data analysis, SB supervised data acquisition, NA carried out data acquisition, NM carried out data acquisition,

KR critically reviewed the drafts. All authors reviewed the draft and assisted in the manuscript preparation.

5.7. Ethics declaration

5.7.1. Competing interests

The authors declare no competing interests.

6. Network analysis-based detection of pathways putatively involved in compensation of drug-resistance-related fitness costs in *Mycobacterium tuberculosis*

Sebastian M. Gygli[1,2,*], Monica R. Ticlla[1,2,*], Andrej Trauner[1,2], Sebastien Gagneux[1,2]

[1]Swiss Tropical and Public Health Institute, Basel, Switzerland
[2]University of Basel, Basel, Switzerland
[*]Authors contributed equally to this chapter

6.1. Abstract

During the past 20-30 years, the epidemic of drug resistant TB has heavily affected the countries of the former Soviet Union and South Africa. Drug resistance was previously assumed to be universally detrimental to *M. tuberculosis* bacteria in the absence of the drug, and drug resistant strains were not thought to be capable of sustaining an infection chain. However, several studies have now demonstrated that the appearance of drug resistant strains is not only due to de novo evolution of resistance within patients, but due to transmission of drug resistant strains. In parallel, studies have reported that drug resistance can come at low or no fitness costs or be accompanied by fitness-restoring compensatory mutations. Thus, the current understanding is that drug resistance in *M. tuberculosis* does not *a priori* exclude transmissibility. However, the demonstrated cases of fitness compensation are largely restricted to reduction of fitness costs due to RIF-resistance due to mutations in a limited set of genes. This is surprising, given the fact that extensively drug resistant *M. tuberculosis* strains are resistant against at least 4 drugs, but frequently carry even more mutations. In the light of complex drug resistance patterns, compensation of drug-resistance related fitness costs might not be driven by mutations in small set of genes, but by a multitude of different mutations affecting e.g. a biological pathway. As the propensity of homoplasy is reduced, as e.g. multiple mutations in different genes of a pathway may confer similar phenotypes, the detection of patterns is dependent on very large, unbiased datasets. Here we apply a network analysis-based approach to detect pathways frequently hit by mutations to identify pathways contributing to successful transmission of MDR *M. tuberculosis* strains in the Republic of Georgia. The preliminary results of these analyses are reported in this working chapter and future directions of research are outlined.

6.2. Introduction

During evolutionary history, obligate pathogens, including *M. tuberculosis*, established and adapted to an ecological niche within their host. Many intracellular pathogens demonstrate low genetic diversity, for reasons not well understood (Achtman, 2008). One possible explanation is "diminishing returns epistasis", wherein most mutations are at best marginally beneficial (Chou *et al.*, 2011) because the long history of coevolution between the human host and *M. tuberculosis* (Brites *et al.*, 2015; Gagneux, 2018) might have led to high levels of adaptation. In this scenario, in early stages of adaptation, a few mutations with large effects on fitness dominate and reach fixation in the population. Given a stable environment, subsequent mutations will find themselves in a genetic background shaped by previous mutations with which they interact epistatically. The number of mutations with large effects diminishes with time and the number of mutations with small effects on fitness increase as a result. While strongly deleterious mutations will vanish due to purifying selection within the host they arose in, beneficial mutations will be prone to extinction due to genetic drift. In these contexts, mutations will have to have a strong selective advantage to reach the numbers within a host required to be transmitted, which, as we have seen previously, is not likely to occur in stable environments and in a well-adapted genetic background. However, there are mutations with large selective benefits in special environments. Mutations conferring drug resistance are the canonical example (Merker *et al.*, 2018). Detection of drug resistance mutations in *M. tuberculosis* is relatively easy, as most high-level resistance is conferred by chromosomal mutations with large selective effects (Coll *et al.*, 2015). However, drug resistance mutations frequently cause reduced growth rates and are only beneficial in the presence of antibiotics. This creates strong selection for mutations restoring (compensating) replicative fitness in absence of the drug (Comas *et al.*, 2012; Merker *et al.*, 2018; Vos *et al.*, 2013). The best known example comes from the compensation of RIF resistance-related fitness costs caused by mutations in the β'-subunit of the DNA-dependent RNA polymerase. Furthermore, mutations in KatG, conferring isoniazid resistance, are known to reduce the catalytic activity of KatG (Yu *et al.*, 2003), which coincides with elevated oxidative stress. These effects may be compensated by upregulation of the peroxiredoxin ahpC (Sherman *et al.*, 1996).

Surprisingly, apart from these two examples, there are no other well-established mechanisms of fitness cost compensation in *M. tuberculosis* caused by chromosomal mutation. The plastic upregulation of TlyA expression in capreomycin resistant mutants, reducing capreomycin resistance-related fitness costs is better described as buffering of fitness costs rather than compensation, as the fitness costs are only partially compensated (Freihofer

et al., 2016). There are multiple possible explanations for this observation. Certain drug resistance mutations might not inflict a fitness deficit (Gagneux *et al.*, 2006b). Furthermore, combinations of drug resistance mutations have been demonstrated to alter the adaptive landscape considerably. Fitness costs of drug resistance mutations may be mitigated by the presence of a further resistance mutation (Borrell *et al.*, 2013). On the other hand, the number of different targets of compensation is greatly enlarged in double mutants (Moura de Sousa *et al.*, 2017). The inflation of potential targets of compensation makes the detection of homoplasic compensatory mutations greatly dependent on the sample size. Furthermore, genetic association studies in bacteria are hampered by extensive linkage disequilibrium, making the separation of loci under selection from the genetic background of a Lineage challenging (Chen *et al.*, 2015). In fact, most tools aiming at detecting homoplasy after adjusting for phylogenetic relatedness, have only been able to detect a small number mutations occurring multiple times (Farhat *et al.*, 2013). The failure to detect homoplasy could be a consequence of the genetic determinants under selection not being reducible to convergence on the mutational or even gene-level, at least not using highly skewed convenience samples as a basis for identifying mutational convergence.

Antibiotics target essential bacterial pathways, and drug resistance mutations can have pleiotropic downstream effects (Perkins *et al.*, 2008). It is therefore plausible that compensation for drug resistance-related fitness costs inflicted by a single mutation, e.g. RpoB S450L for rifampicin (RIF) resistance, may involve multiple different aspects. Recent work on fitness costs of RIF resistance-related fitness costs has demonstrated that different RIF resistance mutations cause different phenotypes, for example deficiencies in initiation, elongation and termination in transcription (Stefan *et al.*, 2018), which in turn might be differentially influenced by compensatory mutations. The different components of fitness costs of RpoB mutations might explain why we see preferential association of certain resistance mutations with secondary mutations in RpoA/C.

Surpassing the demonstrated compensatory effects of single mutations, compensation might also take place at the pathway level. To our knowledge there is currently no known example of convergent evolution on the pathway level compensating for drug resistance-related fitness costs. However, mutations targeting the gene alr in *M. tuberculosis* confer resistance to cycloserine, by abrogating the utilisation of pyruvate. The downstream effect of loss of function mutations in alr is that there is an overabundance of the substrate of the drug target, inhibited by cycloserine (Desjardins *et al.*, 2016b). In principle, such a process would also be a possible mechanism for compensating drug resistance-related

fitness costs resulting from drug resistance mutations in essential enzymes, for example those that lead to lower enzymatic processivity. If inactivating any enzyme in a metabolic pathway would have the same downstream effect, e.g. disrupting the possibility to utilize pyruvate, there would be little convergence expected on the mutation or gene level, as any disruptive mutation in any of the genes in the pathway would have the same effect.

Here, we report a first attempt to discover pathways under selection in multidrug-resistant *M. tuberculosis* strains, which could be involved in compensation of drug resistance-related fitness costs. To do this, we applied a network analysis-based approach developed by Verbeke *et al.* (Verbeke *et al.*, 2015), which uses prior information derived from gene-gene interaction networks together with mutation data to infer which pathways are affected more frequently by mutations than expected by chance. As an input, we used mutational data extracted from the population-based *M. tuberculosis* sample set described in chapter 5. We first identified strains belonging to transmission clusters and then extracted mutations specific to the transmission cluster. These mutations putatively contribute to the success (transmissibility) of these strains.

6.3. Materials and Methods

6.3.1. *M. tuberculosis* strains & transmission cluster inference

For the analyses described in this chapter, we used the mutational data generated in chapter 5 as an input. Briefly, we used a sample set of 659 clinical M. tuberculosis strains displaying at least an MDR phenotype, isolated in Georgia between 2011 – 2013 and subjected them to whole genome sequencing. The sample set consists of 53 % of reported MDR *M. tuberculosis* strains in the sampling timeframe of the study. We used the mutational data generated from whole genome sequences to create a variable position pseudoalignment which served as a basis for inferring a distance matrix. To identify transmission clusters based on genetic distance, we applied hierarchical agglomerative clustering with the unweighted pair-group average method implemented in the agnes function of the R package cluster (v.2.0.6.) to the previously inferred distance matrix. We chose to cut the resulting tree at an average of 5 single nucleotide variants, as this has proven to be a sensible cut-off for likely patient-to-patient transmission (Walker *et al.*, 2013).

6.3.2. Inferring mutations private to transmission clusters

One major issue with genome-wide association studies in bacteria is linkage disequilibrium. This is especially problematic for *M. tuberculosis*, which does not undergo horizontal gene exchange, leading to a clonal population structure. One way to address this problem is by using population-based sample sets. Although we only sequenced ca. 53 % of all MDR *M. tuberculosis* cases isolated in Georgia between 2011 and 2013, the sample set is not biased towards outbreaks. To extract mutations private to transmission clusters, we compared the drug resistance profile of transmission clusters to the drug resistance profile of their parent nodes. Parent nodes sharing the same drug resistance profile as a given cluster were assumed to belong to the same, larger transmission cluster and that mutations compensating for fitness costs were acquired earlier. We therefore compared drug resistance profiles from a given cluster to all of its parent nodes until the drug resistance profile differed by more than 2 mutations. We then took the set difference between the intersect of all mutations in a cluster with the parent node, resulting in the mutations private to the transmission cluster. We then filtered the list for drug resistance mutations acquired since divergence from parent.

6.3.3. Pathway analysis

We used all mutations which were private to the different transmission clusters to rank pathways by relevance using an adaptation of Verbeke *et al.*'s network-based approach. Although the original approach integrates multiple types of data (gene expression, mutation, methylation and copy number data) together with *a priori* knowledge about gene-gene interactions to rank pathways, we focused on using only mutation data in this first attempt. Our adaptation constructs a global network where *M. tuberculosis* strains, mutations and genes are represented as nodes. Relationships between nodes, represented as undirected edges, include: i) the presence of a mutation in a particular strain, ii) the assignment of a mutation to a particular gene (intergenic mutations were assigned to a particular gene if the mutation located before the start codon) and iii) the physical or functional interaction between genes. We constructed a reference gene-gene interaction network derived from KEGG pathways; pathways only containing a single gene were discarded. As described by Verbeke *et al.*, the degree of relatedness between each pair of nodes was inferred by applying the Laplacian Exponential Diffusion (LED) kernel on the graph, followed by normalization, producing a global similarity matrix. Since applying a kernel on graph nodes assumes a fully connected network, we applied the kernel on the largest connected component of our global network if it was not fully connected. The LED kernel is in-

fluenced by a diffusion parameter (α), which was set to 0.1; however, results were stable between 0.001 and 0.1. Given an *M. tuberculosis* strain and a pathway composed of a set of genes, the average of the similarities between the *M. tuberculosis* strain and each gene in the pathway corresponds to the pathway importance score for that *M. tuberculosis* strain. The pathway importance score for a set of *M. tuberculosis* strains is obtained by averaging their importance scores. To get the probability of obtaining an importance score at least as extreme as the one obtained by chance, a random score distribution is obtained for the pathway. This random distribution is obtained by repeatedly calculating the pathway importance score from sets of genes randomly sampled from the similarity matrix. The size of the randomly sampled gene sets is equal to the size of the pathway gene set. Thus the pathway importance score is ultimately converted to a p-value.

6.4. Results and Discussion

6.4.1. Mutations private to transmission clusters

As reported in chapter 5, there were a total of 42 transmission clusters with a median size of 3.5 strains (figure 6.1) and a total of 301 clustered strains. We identified a total of 1192 mutations private to transmission clusters out of a total of n = 10435 variable positions in the complete dataset, in a total of 1372 genes. Most mutations were only shared by 1-2 clusters, indicating that the approach is successful in removing most of the phylogenetic signal. (Figure 6.2). A number of SNVs do, however, occur in a large proportion of strains, indicating the presence of phylogenetic markers. To avoid including phylogenetic markers, the "tree walking" algorithm could be altered. Instead of looking for parent nodes of clusters with differing drug resistance profiles, the algorithm could be purely distance-based. From the work of Walker *et al.*, we know that the likelihood of direct transmission between two patients drops if the pairwise genetic distance between the two patient isolates is larger than 5 mutations. This could be exploited by including parent nodes until the genetic distance between the two patient isolates within a cluster reaches a certain threshold. The threshold could be initially set at 12 mutations, as Walker *et al.* demonstrated that when strains differed by more than 12 mutations, no epidemiological links between two strains could be identified (Walker *et al.*, 2013). The mutations private to the transmission cluster would then again be determined by the set difference between the intersect of all mutations within the transmission cluster and the parent node.

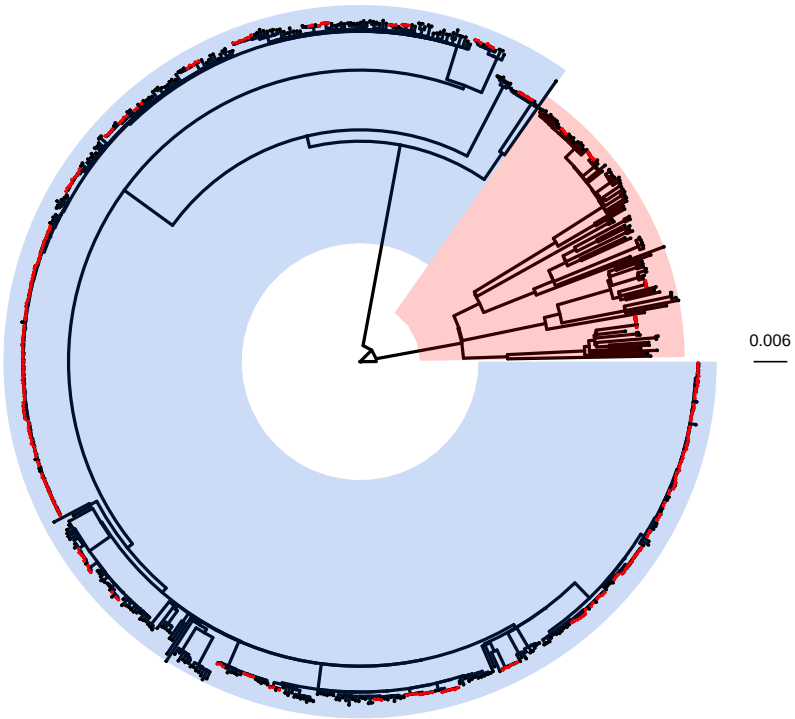

Figure 6.1.: **Maximum likelihood phylogeny of the 659 *M. tuberculosis* strains included in the network analysis study.** Strains belonging to clusters inferred by hierarchical agglomerative clustering are colored in red. Blue shading indicates strains belonging to Lineage 2, red shading indicates strains belonging to Lineage 4. Scale bar indicates substitutions per site with a total of 10435 variable positions.

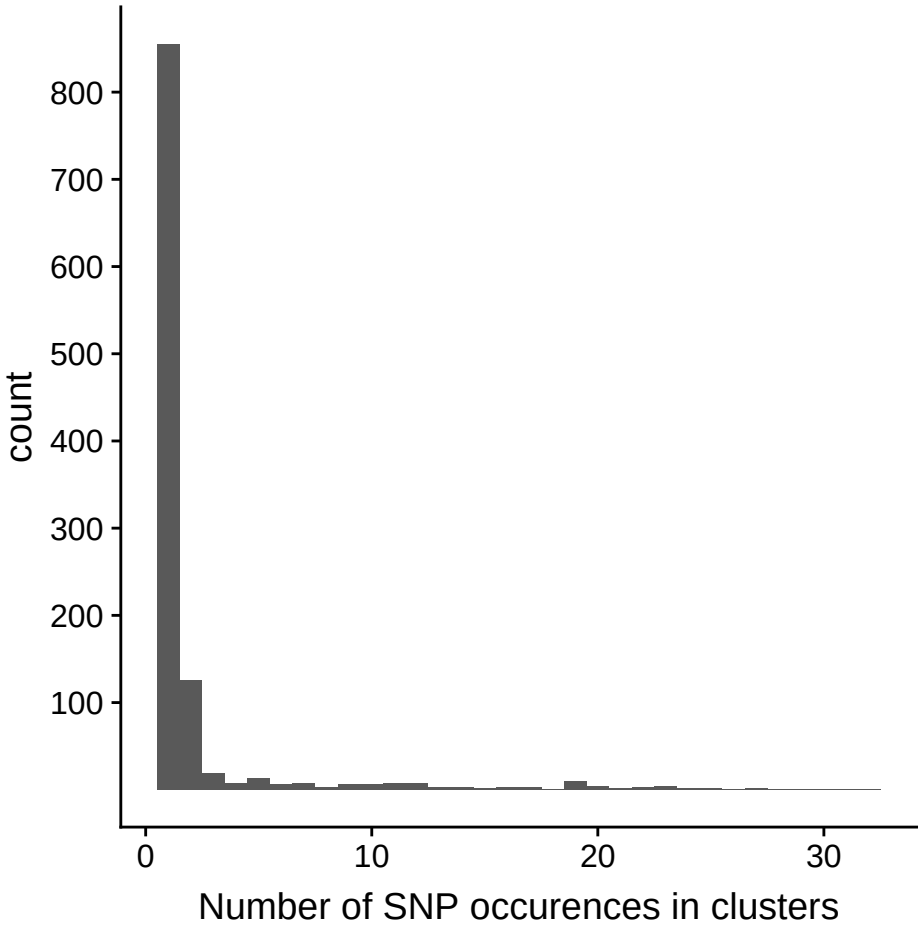

Figure 6.2.: **Histogram of how frequently mutations occurred in clusters.** Certain mutations appeared in 30 clusters and were most likely phylogenetic markers.

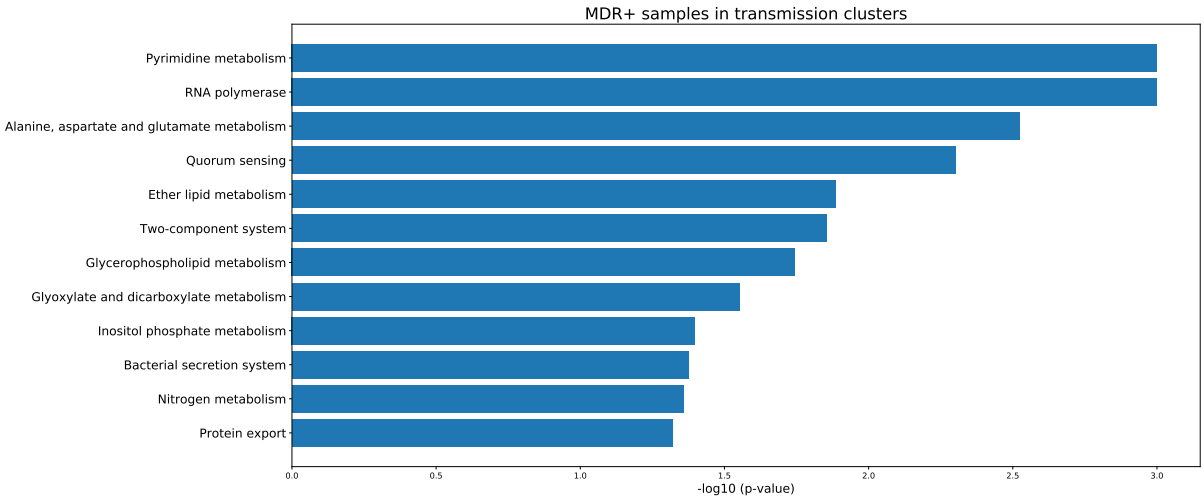

Figure 6.3.: **Pathways with the highest score.**

6.4.2. Pathway relevance ranking– proof of concept and intriguing findings

Figure 6.3 shows the pathways that carried mutations more often than expected by chance. A total of 12 pathways were hit more frequently than expected by chance at a significance level of $p < 0.05$. Most pathways show hits in multiple genes, whereas most individual genes do not show multiple mutations (figure 6.4). The RNA polymerase pathway was among the highest ranked pathways; the genes most frequently affected in those pathways were *rpoA* and *rpoC* (figure 6.4). This served as a proof of concept, as we expected the pathways containing the known targets of compensation to be enriched. However, it is interesting that the pyrimidine biosynthesis pathway, as well as alanine, aspartate and glutamate metabolism (nucleotide precursors) pathways are ranked highly, as this might hint at a connection to the RNA polymerase and therefore RIF resistance. We observed a total of 64 unique mutations in the pathways involved in nucleotide biosynthesis. Excluding mutations in the RNA polymerase, RIF-resistant strains have more non-synonymous mutations in genes belonging to the pyrimidine and nucleotide precursor synthesis compared to RIF-susceptible strains (median number of mutations in RIF resistant = 35 vs. susceptible = 30, $p < 2.2 \times 10^{-16}$, Mann-Whitney U test). The RIF-resistant RNA polymerase might have altered affinities for nucleotides, potentially requiring more nucleotides for optimal function. Indeed, work performed on RIF-resistant *Escherichia coli* demonstrated that RIF-resistant RNA polymerases have a lower affinity for nucleotides (Jin *et al.*, 1991). This raises the question of whether mutations in nucleotide biosynthesis pathways are involved in fitness-compensation, independent of the effect that *rpoC* mutations have on restoring transcription in *rpoB* mutants (Stefan *et al.*, 2018). Furthermore, preliminary metabolomics data generated from 3 clustered and 3 unclustered strains picked from the same dataset used in this analysis, clustered strains show increased uracil concentrations compared to unclustered strains. The clustered strains share two nonsynonymous mutations in carB (pyrimidine biosynthesis pathway) and apt (purine biosynthesis pathway). Furthermore, the Georgian dataset revealed that, although many canonical compensatory mutations have been detected (e.g. RpoA T184A, RpoC V483G), there were some that locate to the secondary channel of the RNA polymerase and may interact with nucleotides. These results hint at the potential existence of alternative pathway for compensating RIF-resistance related fitness costs in *M. tuberculosis* either via altering the RNA polymerase, potentially to increase the affinity of the RIF resistant enzyme to nucleotides or by increasing nucleotide levels.

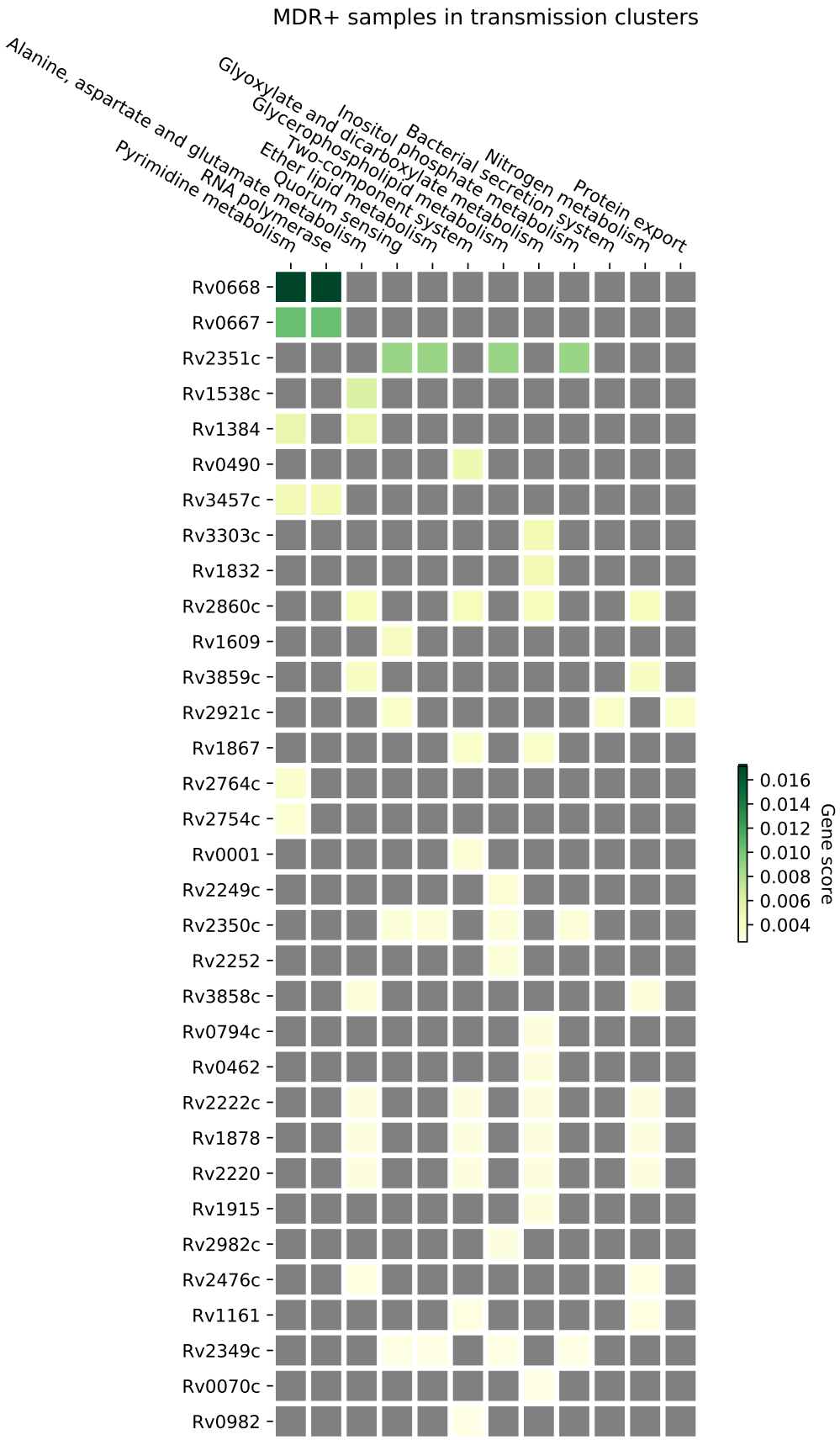

Figure 6.4.: **Genes within the pathways with the highest relevance scores.** Genes are color coded according to the number of different mutations detected within a gene.

There are is a caveat in that the wealth of mutations detected in the nucleotide biosynthesis pathways spread over a number of different genes does not offer any directly apparent mechanism in which non-synonymous mutations in these genes could increase nucleotide levels. However, the regulation of nucleotide biosynthesis in bacteria is complex in general (Turnbough *et al.*, 2008), and there is little information on regulation of these pathways in *M. tuberculosis*. More detailed metabolomics experiments of strains carrying these mutations will be required to link genotypes and phenotypes.

6.4.3. Other highly ranked pathways and potential connections

Mutations identified in the quorum sensing pathway mainly map to the gene Rv2351c, which encodes a phospholipase. However, mapping the mutation on the phylogenetic tree revealed that all Lineage 2 strains have acquired 3 mutations within this gene, confirming the mutation list still contains phylogenetic variants. The high ranking of pathways associated with lipid metabolism (glycerophospholipid & ether lipid metabolism), as well as inositol phosphate metabolism might indicate a connection to cell wall biosynthesis. Lipids are a major constituent of the mycobacterial cell (Jackson, 2014) wall and inositol phosphate is crucial to link arabinogalactan to the outer cell membrane (Movahedzadeh *et al.*, 2010). Multiple antituberculous drugs, including including isoniazid, ethambutol, pyrazinamide and ethionamide, target cell wall biosynthesis. The contribution of two-component systems to compensation of fitness costs is more difficult to establish. Two-component systems perform a multitude of different tasks in intracellular signalling and are grouped in this pathway based on their general function (sensor kinases), rather than acting in a concerted manner.

6.5. Future directions

As a first step, we propose to repeat the identification of mutations private to transmission clusters purely on a distance-based measure to be able to exclude the vast majority of phylogenetic mutations. The detection of a potential new pathway of compensation is exciting, and it is not surprising that it is related to RIF resistance as resistance to this drug is known to inflict substantial fitness deficits in some strains. However, the method also detected several pathways related to cell wall biosynthesis. We would like to investigate these pathways in more detail by stratifying the dataset into groups of strains

resistant to drugs targeting cell wall biosynthesis vs. others. For this, we need to include more samples to increase numbers in the different compartments. Sequencing efforts in Georgia continue and a wealth of strains from subsequent years (2013 onwards) will be available soon.

6.5.1. Hypotheses

We furthermore want to follow up on the potential involvement of nucleotide levels in the compensation of RIF-resistance related fitness costs concerning the following hypotheses:

1. Fitness costs of RIF resistance mutants vary as a function of pyrimidine (uracil) availability.
2. Mutations in the secondary channel of the RNA polymerase influence nucleotide affinity and reduce fitness costs of RIF resistance.
3. Mutations found in nucleotide biosynthesis pathways do not occur or occur less frequently with RNA polymerase mutations in the secondary channel.
4. Strains with mutations in nucleotide precursor or pyrimidine biosynthesis pathways demonstrate higher levels of uracil.

6.5.2. Objectives

We want to test these hypotheses by addressing the following objectives

1. Generate pyrimidine auxotrophic Mycobacterium smegmatis strains by disrupting the pyr operon responsible for pyrimidine biosynthesis in RIF resistant backgrounds involving a high (RpoB H445Y) and low cost (RpoB S450L) mutation. Perform growth curves in presence of varying uracil concentrations.
2. In the backgrounds outlined under point 1, generate strains carrying mutations in the secondary channel of the RNA polymerase and perform growth curves in varying uracil concentrations.
3. Using the Georgian MDR *M. tuberculosis* strain set from 2011-2013 and subsequent years, assess patterns of coexistence/mutual exclusivity of mutations in nucleotide biosynthesis pathways and mutations in the secondary channel of the RNA polymerase.

4. Perform large scale metabolomics analyses of clustered (n = 150) and unclustered (n = 100) MDR *M. tuberculosis* strains and assess differences between the two groups with a special emphasis on nucleotide levels (project is already underway).

6.5.3. Expected outcomes

Reduced pyrimidine levels in RIF resistant backgrounds should inflate the fitness costs, as the affinity of the RNA polymerase for pyrimidines is reduced compared to RIF susceptible strains. Furthermore, mutations in the secondary channel of the RNA polymerase should increase the growth rate of RIF resistant strains dependent on the uracil concentrations in the medium. If there are two ways of coping with the lower affinity of RIF resistant RNA polymerases, namely increasing the intracellular pyrimidine levels or restoring affinity via mutations in the secondary channel of the RNA polymerase, the two pathways should be, to some degree, mutually exclusive. We expect that strains we inferred as clustered will differ from unclustered strains in terms of intracellular nucleotide levels; clustered strains would be predicted to have higher pyrimidine levels. If this effect is observed, the large scale of the experiment will help to establish a genotype-phenotype relationship between mutations in the nucleotide precursor and biosynthesis pathways and nucleotide levels.

7. Evolution of drug resistance under sub-inhibitory drug concentrations alone or in combination in *Mycobacterium tuberculosis*

Sebastian M. Gygli[1,2], Miriam Reinhard[1,2], Sonia Borrell[1,2], Sebastien Gagneux[1,2]

[1]Swiss Tropical and Public Health Institute, Basel, Switzerland
[2]University of Basel, Basel, Switzerland

7.1. Abstract

Drug-resistant *M. tuberculosis* variants may evolve in single patients even under optimal treatment conditions. It is believed that sub inhibitory drug concentrations (sub-MIC) contribute to treatment failure, by giving resistant variants, omnipresent in populations of sufficient size, a selective benefit over drug-susceptible variants. Drug efflux mechanisms have been implicated in low-level drug resistance and their induction by the presence of antituberculous drugs has been demonstrated. *M. tuberculosis* infections are treated with at least 4 active antituberculous drugs- However, the influence of drug combinations at sub-MIC on the evolution of drug-resistance in *M. tuberculosis* is unknown. Here we present a study investigating the influence of sub-MIC of two drugs, rifampicin (RIF) and isoniazid (INH), alone and in combination on the evolution of drug resistance in three drug-susceptible clinical *M. tuberculosis* isolates. The strains were passaged for ≈180 bacterial generations in the presence increasing concentrations of RIF and INH alone and in combination, starting at a fraction of the wt MIC and ending at 4 x MIC. Whole genome sequencing was performed of three endpoint cultures per treatment and strain. We report the detection of drug-resistance conferring mutations in most cultures that survived until the end of the experiment.

7.2. Introduction

The *de novo* evolution of drug-resistant *M. tuberculosis* variants is generally thought to be a result of inadequate treatment due to interrupted drug supply, delayed drug susceptibility testing, prescription of inadequate treatment regimens, patient non-adherence, or low drug quality (Law *et al.*, 2017). Furthermore, there are a number of factors specific to human hosts which can generate sub-MIC drug environments within patients, also resulting in selection for resistant variants (Gullberg *et al.*, 2011). Patients may differ in terms of how well drugs are absorbed and how fast they are metabolized and rendered ineffective, as demonstrated for RIF and INH (Gumbo *et al.*, 2007; Van Ingen *et al.*, 2011; Pasipanodya *et al.*, 2013). Furthermore, antituberculous drugs have vastly different chemical properties and as a result, drugs differ in their capacity to diffuse into mycobacteria-containing tissues. Unequal penetration of drugs in combination therapy will generate spatio-temporal variation in drug concentrations (Prideaux *et al.*, 2015) known to select for multidrug-resistance (MDR: resistance against rifampicin and isoniazid) (Moreno-Gamez *et al.*, 2015). Gullberg *et al.* demonstrated that sub-MIC drug concentrations lead to the stable coexistence of resistant and susceptible variants, or can even select for a fully resistant population. Drugs at sub-MIC will reduce the growth rate of susceptible strains in a dose dependent manner, creating an ecological niche for drug-resistant bacteria suffering from drug-resistance-related fitness costs. If the growth rate of the drug-susceptible strains is reduced to the level of the resistant variants, coexistence of resistant and susceptible variants in the same population is the result. Drug-resistant variants will outgrow the drug-susceptible strains if the growth rate of the drug-susceptible strains is reduced to levels below that of the resistant variant, or if the drug resistant variants acquire compensatory mutations. The lowest drug concentration at which resistant strains have a selective benefit is termed the "minimal selective concentration" (MSC). There are no data on the MSC for antituberculous drugs alone or in combination.

Regimens containing multiple drugs are highly effective in treating TB infections and preventing the evolution of resistant variants in *M. tuberculosis* (Feng-Zeng *et al.*, 1996). Yet, even under optimal conditions, drug resistant variants may evolve in individual patients (Bloemberg *et al.*, 2015; Srivastava *et al.*, 2011). It has been hypothesized, that the evolution of drug resistant variants in *M. tuberculosis* is due to "functional monotherapy" as a result of one of the factors described above that reduces drug concentrations to sub-MIC levels (Colijn *et al.*, 2011; Warner *et al.*, 2006). M tuberculosis strains may also find themselves in the presence of multiple drugs at sub-MIC levels. The influence of sub-MIC drug combinations on the evolution of MDR *M. tuberculosis* strains is largely unexplored

(Ankomah *et al.*, 2012). Apart from selecting for chromosomal resistance, it is known that sub-MIC drug levels induce the expression of drug efflux pumps (Gygli *et al.*, 2017), potentially serving as a stepping stone for high-level, chromosomally encoded resistance. The influence of efflux pumps on generating high-level resistance is a matter of current debate. There seems to be great variability in the expression of efflux pumps among clinical isolates (Li *et al.*, 2015; Machado *et al.*, 2017). However, most strains displaying clinically relevant levels of RIF and/or INH resistance have chromosomal mutations in *rpoB* and *katG*, respectively. In the clinics, RIF resistance is used as a surrogate marker for MDR strains, as RIF mono-resistance is rare. In contrast, INH mono-resistance is frequently observed.

Here we report preliminary analyses of a serial passage experiment involving three clinical *M. tuberculosis* strains exposed to increasing concentrations of RIF and INH alone and in combination, starting at a fraction of the wild type MIC and ending at 4 x MIC. Whole genome sequencing was performed on three endpoint replicates (endpoint: 180 generations) per strain and treatment.

7.3. Materials and methods

7.3.1. *M. tuberculosis* strains

We selected three clinical *M. tuberculosis* isolates from three different genetic backgrounds: a Lineage 2 strain (N0155, Beijing sublineage) due to the frequent association of Lineage 2 with drug resistance phenotypes (Chapter 3), a Lineage 4 strain (N1283, Ural sublineage), as Lineage 4 strains are endemic in Georgia, the site of a large sampling project (see Chapters 5 and 6), in addition to being frequently associated with drug resistance, and a Lineage 1 strain (N0157, Manila sublineage), a Lineage in which drug resistance is relatively rare, despite causing a considerable number of cases in South-East Asia (Roa *et al.*, 2018).

To select individual colonies, strains were plated on Middlebrook 7H10 agar and incubated at 37° C for 3 weeks. 10 individual colonies were picked and sub-cultured in Middlebrook 7H9 medium to an OD_{600} of ≈0.5. Stocks equivalent to $OD_{600} = 0.5$ were generated by spinning the cells down, removing the culture supernatant and resuspending the pellet in the respective amount of H_2O containing 10 % (v/v) glycerol.

7.3.2. Culture medium & culture vessel

For all experiments, we applied Middlebrook 7H9 ADC liquid medium, supplemented with 0.1 % Tween80 to reduce clumping. Serial transfer experiments were performed in 10 ml standing cultures in inkwells sealed with a vented cap to allow for air exchange. Biosafety concerns due to the likely generation of MDR strains made the use of agitated cultures unfeasible. Growth properties in standing cultures were assessed by measuring the optical densities achieved after 21 days of incubation of 6 replicate cultures per strain. Two different inocula were evaluated: 5 x 10^6 cells and 0.5 x 10^6 cells.

7.3.3. MIC determination

We determined the MICs for the selected strains for RIF, INH and RIF and INH in combination using a resarzurin microtiter assay (Franzblau *et al.*, 2012) in a checkerboard layout (Orhan *et al.*, 2005). We screened the strains for susceptibility to drug concentrations in twofold dilutions. For RIF, the concentration range assayed was 6.25 x 10^{-3} μg/ml – 2.44 x 10^{-5} μg/ml and for INH, 40 μg/ml – 1.56 x 10^{-3} μg/ml. For the screening, we cultured the strains to mid log phase (OD_{600} ≈0.5), diluted the cultures to 10^6 cells/ml and added 10 μl of the diluted suspension to the microtiter wells containing 90 μl Middlebrook 7H9 medium including the drugs. We incubated the plate for 6 days at 37° C, after which 10 μl 0.02 % (w/v) resazurin was added to the wells and incubated overnight. A fluorometer was used to measure the concentration of metabolized resazurin as an indicator of bacterial metabolic activity (excitation wavelength 572 nm, emission wavelength 583 nm). "No bacterial growth" at a given drug concentration was defined as a reduction of the relative fluorescence by > 90 % compared to a no-drug control.

7.3.4. Serial passage

There were 10 replicate cultures per genetic background and treatment. There were in total four treatments: RIF alone, INH alone and RIF + INH, as well as a no-drug control. As there is no prior information on the MSC for any drug in *M. tuberculosis*, we decided to start the serial passage experiment at 1/64 of the MIC. Initial cultures were set up by inoculating 20 μl of the calibrated stock suspension ($OD_{600} = 0.5$) into 10 ml of 7H9 ADC medium containing 0.1 % tween 80. Due to the light sensitive nature of RIF, cultures of treatment 1 and 3 were protected from light. Cultures were grown for 3 weeks at 37° C and sub cultured by inoculating 10 μl of the grown culture into 10 ml of fresh medium

(see results). Every second passage, stocks were frozen and the drug concentration was doubled.

7.3.5. Fractional inhibitory concentration index (FIC)

The FIC for a drug A was calculated by dividing the MIC of drug A in combination with drug B by the MIC of drug A alone. The sum of the FIC of drug A and drug B determined the interaction between the two drugs: σFIC $\leq$ 0.5: synergistic, σFIC $>$ 0.5 $<$ 2: indifferent, σFIC $\geq$ 2: antagonistic (Orhan *et al.*, 2005).

7.3.6. Growth indices

All cultures appeared to grow robustly until generation 110 (passage 11 – 0.5 x MIC). Growth indices were collected from generation 110 onward by assessing robustness of growth by eye using the following keys: 1: scant growth, 2: intermediate growth, 3: robust growth.

7.3.7. Whole genome sequencing

We sequenced three end point cultures from generation 180 from treatments 1, 2, 3 (only N0155 and N1283) and 4. As no replicates of N0157 survived until the end of the experiment in treatment 3, we sequenced 3 replicates from generation 130 cultures instead. Whole genome sequencing, mutation identification and annotation was performed as described previously (Ghielmetti *et al.*, 2017; Gygli *et al.*, 2018).

7.4. Results

7.4.1. MIC for RIF, INH and RIF + INH combinations

All strains demonstrated similar MIC values for the two drugs alone and in combination (table 1) and differed by at most one dilution step. The interaction between INH and RIF in combination was indifferent for N1283 (σFIC = 2) and N0157 (σFIC = 2) and antagonistic for N0155 (σFIC = 3). Indifference between RIF and INH has been reported previously (Dickinson *et al.*, 1977). As the maximum difference in MIC between the strains was within one dilution step, we used the same MIC for all strains to reduce the complexity of the experiment.

Table 7.1.: **MICs for isoniazid and rifampicin alone and in combination.**

Strain	MIC INH only	MIC INH in combination
N0155	0.05 μg/ml	0.05 μg/ml
N0157	0.05 μg/ml	0.05 μg/ml
N1283	0.05 μg/ml	0.1 μg/ml
	MIC RIF only	**MIC RIF in combination**
N0155	1.563 ng/ml	1.563 ng/ml
N0157	0.781 ng/ml	0.781 ng/ml
N1283	0.781 ng/ml	0.781 ng/ml

7.4.2. Assessment of growth in standing cultures

M. tuberculosis achieves the highest growth rates in well-oxygenated cultures. However, due to biosafety concerns in selecting for MDR *M. tuberculosis* variants, cultures were not subjected to continuous shaking. To test growth properties of the three selected strains, we assessed growth in 10 ml standing cultures with vented caps. All strains achieved similar optical densities after 3 weeks of incubation with both inocula (OD_{600} = 0.8-1.1). We chose to subculture every three weeks applying a bottleneck of 1:1000, corresponding to ca. 0.5 x 10^6 cells, to achieve ≈200 generations within a manageable time frame (1 – 1.5 years) (Comas *et al.*, 2012; Reynolds, 2000).

7.4.3. Whole genome sequencing results

The median fold coverage of the sequenced strains was 68.0x (25^{th} percentile = 52.7x, 75^{th} percentile = 78.6x, interquartile range (IQR) = 25.9).The median mapping percentage was 98.9 % (25^{th} percentile = 98.8 %, 75^{th} percentile 99.0 %, IQR = 0.2).

Mutations detected in the ancestors and the strains from treatment 4 (no drug control) were removed from the mutations detected in strains from treatment 1-3, to exclude putative media adaptations and to generate a list of mutations which were acquired during serial passage. There were a median of 2 fixed (25^{th} percentile = 1, 75^{th} percentile = 4, IQR = 3), 23 unfixed (25^{th} percentile = 17.5, 75^{th} percentile = 28, IQR = 10.5) and 3 unfixed indels (25^{th} percentile = 2, 75^{th} percentile = 4, IQR = 2) detected per culture.

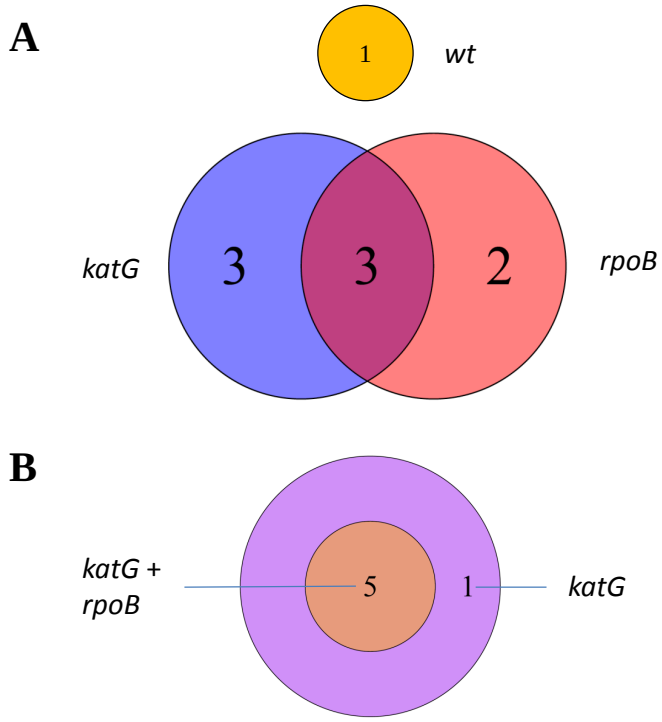

Figure 7.1.: **Mutations detected in canonical resistance genes.** **A** Mutations in drug resistance genes in strains from treatment 1 (RIF). **B** Mutations in drug resistance genes in strains from treatment 3 (RIF + INH).

Table 7.2.: ***rpoB* and *katG*** mutations detected in the three treatments

Target gene	Treatment 1 (RIF)	Treatment 2 (INH)	Treatment 3 (RIF + INH)
rpoB	Ser450Leu, Asp435Gly, Leu430Arg, Asp435Ala	-	Gln172, Ser450Trp, Asp435Gly, Ser441Leu, Val170Phe
katG	Gly285Asp, Gly120Asp Met105Ile, Val1?, Leu415Met, Lys274Glu	Trp328Arg , Val1?, Gly699Trp, Ser315Asn, Asn138Ser, Trp438Arg, Tyr426*, Ala172Thr, Tyr678*	Thr322Ala, Arg484His, Ser315Arg, Trp198*, His270Tyr, Trp341Leu, Gly124fs

7.4.4. Mutations in genes encoding efflux pumps

As a first step, we were interested in seeing if there were any mutations in regulatory regions of known/putative efflux pumps (Silva *et al.*, 2016), potentially conferring low level resistance to RIF and/or INH. We were unable to detect mutations in genes previously associated with drug efflux, potentially indicating that overexpression of efflux pumps is not a relevant mechanism for RIF and/or INH resistance.

7.4.5. Mutations in drug resistance conferring genes

We observed mutations in canonical RIF and INH resistance-related genes (*rpoB* and *katG*). No mutations were observed in the gene *inhA* (target of activated INH) or its promoter. Surprisingly, we identified *katG* mutations in treatment 1 (RIF), in absence of any *rpoB* mutations, however not the opposite (no *rpoB* mutations in treatment 2 (INH)) (Figure 7.1). In total, six out of nine sequenced strains from treatment 1 carry *katG* mutations. Furthermore, two out of three N1283 replicates from treatment 1 do not carry any *rpoB* mutations, yet survive at 4x of the RIF MIC. All strains in treatment 2 harboured *katG* mutations. In treatment 3, all but one strain of the replicates surviving until generation 180 harboured *katG* and *rpoB* mutations. Of the three replicates of N0157 from generation 130, two harboured frame shift mutations in *katG*, but no *rpoB* mutations.

Table 7.3.: **Proportion of replicates surviving until generation 180.**

Strain	Treatment 1 (RIF)	Treatment 2 (INH)	Treatment 3 (RIF + INH)
N0155	100 %	100 %	30 %
N0157	100 %	100 %	0 %
N1283	60 %	100 %	70 %

7.4.6. Growth indices, survival and extinction

In all treatments, strains demonstrated robust growth until generation 110 (0.5 x MIC), after which growth inhibition was observed in a number of cultures. We observed fluctuating growth indices in all treatments as a function of increasing drug concentrations (figures 7.2, 7.3, 7.4). Growth indices decreased when the antibiotic concentration was doubled and increased again after the subsequent dilution step where the strains spent another 3 weeks at the same antibiotic concentration. The different strains demonstrated different mean growth indices within the three treatments, indicating that there are

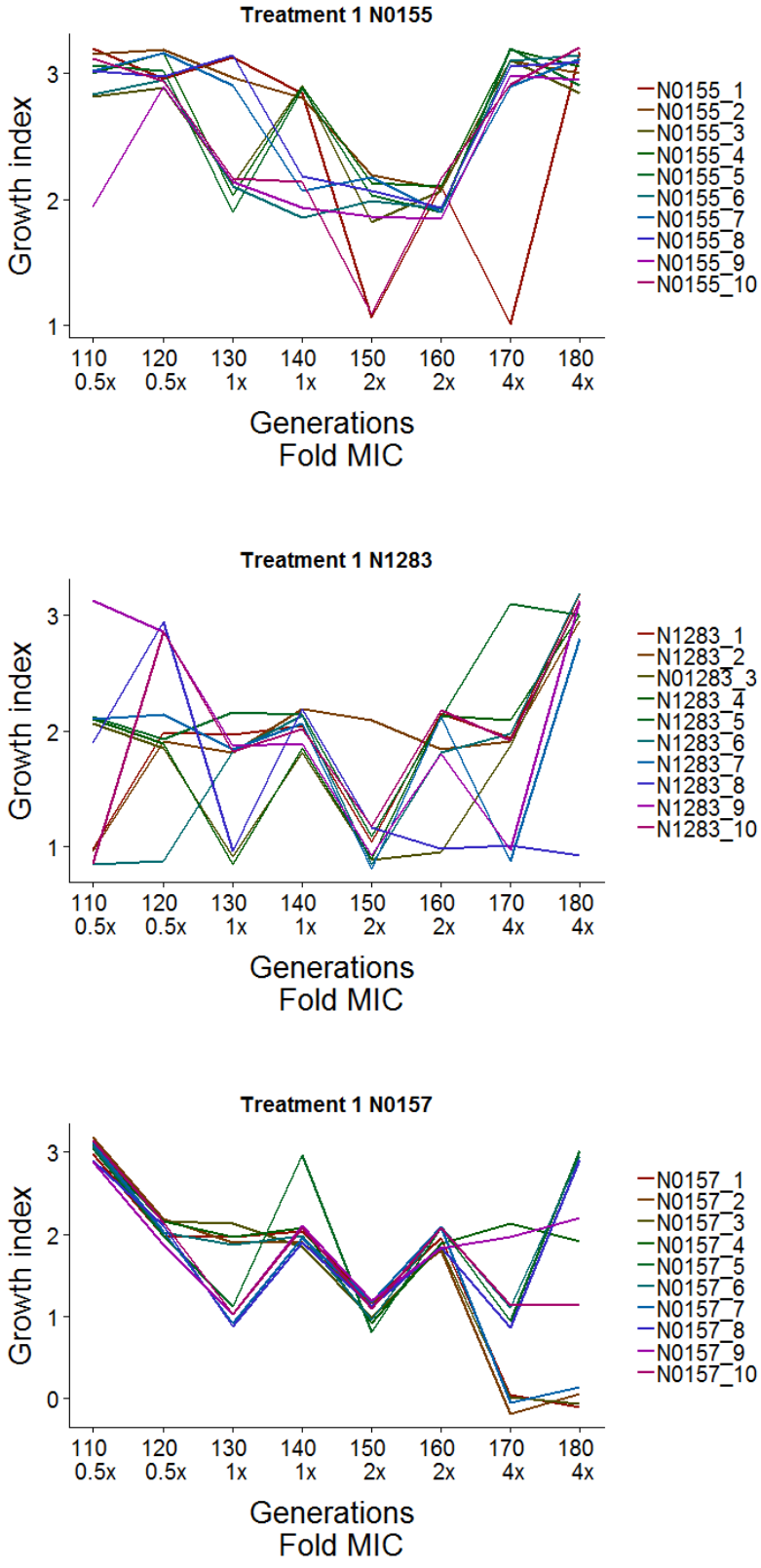

Figure 7.2.: **Growth indices of all cultures in treatment 1 (RIF) starting at generation 110 (0.5 x MIC).**

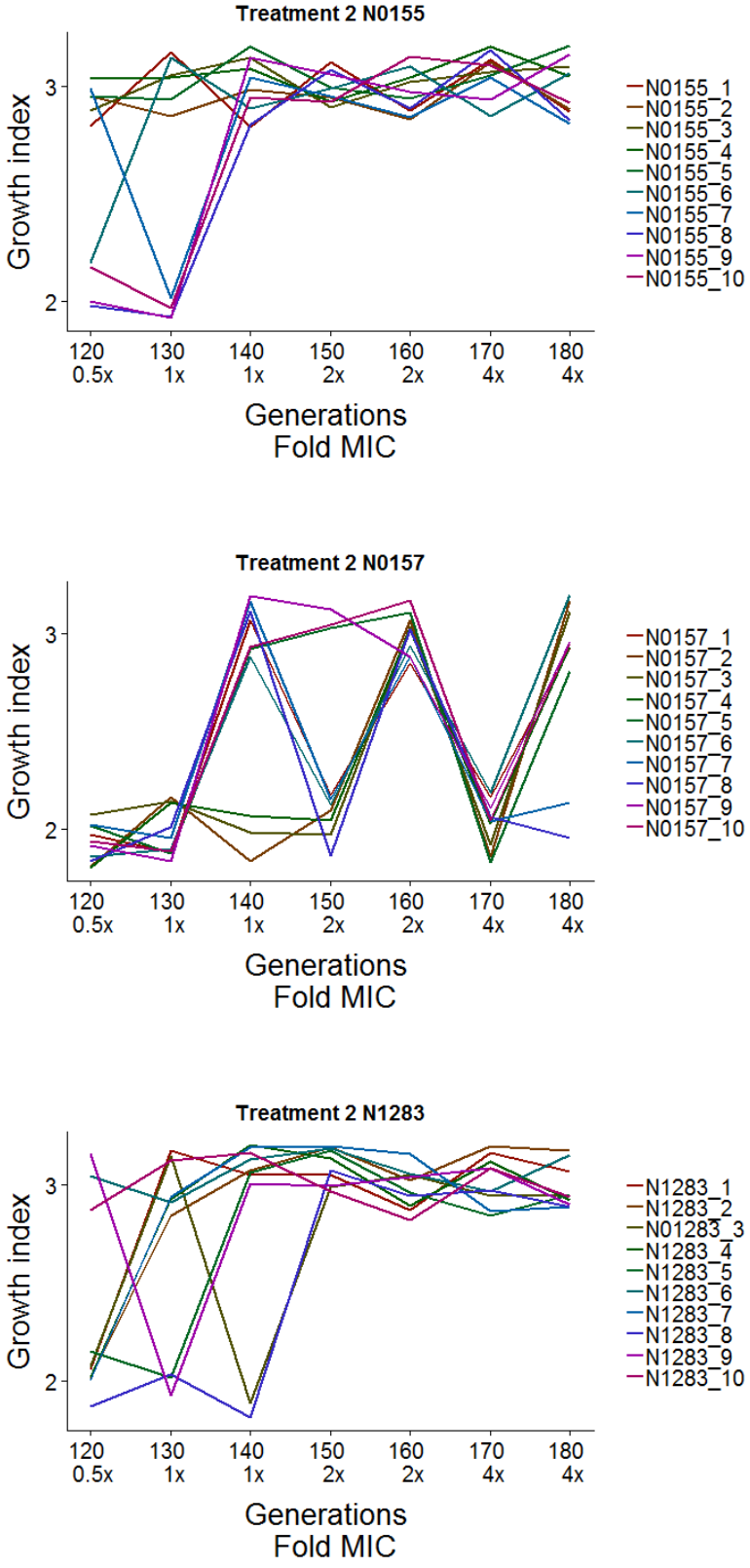

Figure 7.3.: **Growth indices of all cultures in treatment 2 (INH) starting from generation 120 (0.5 x MIC).**

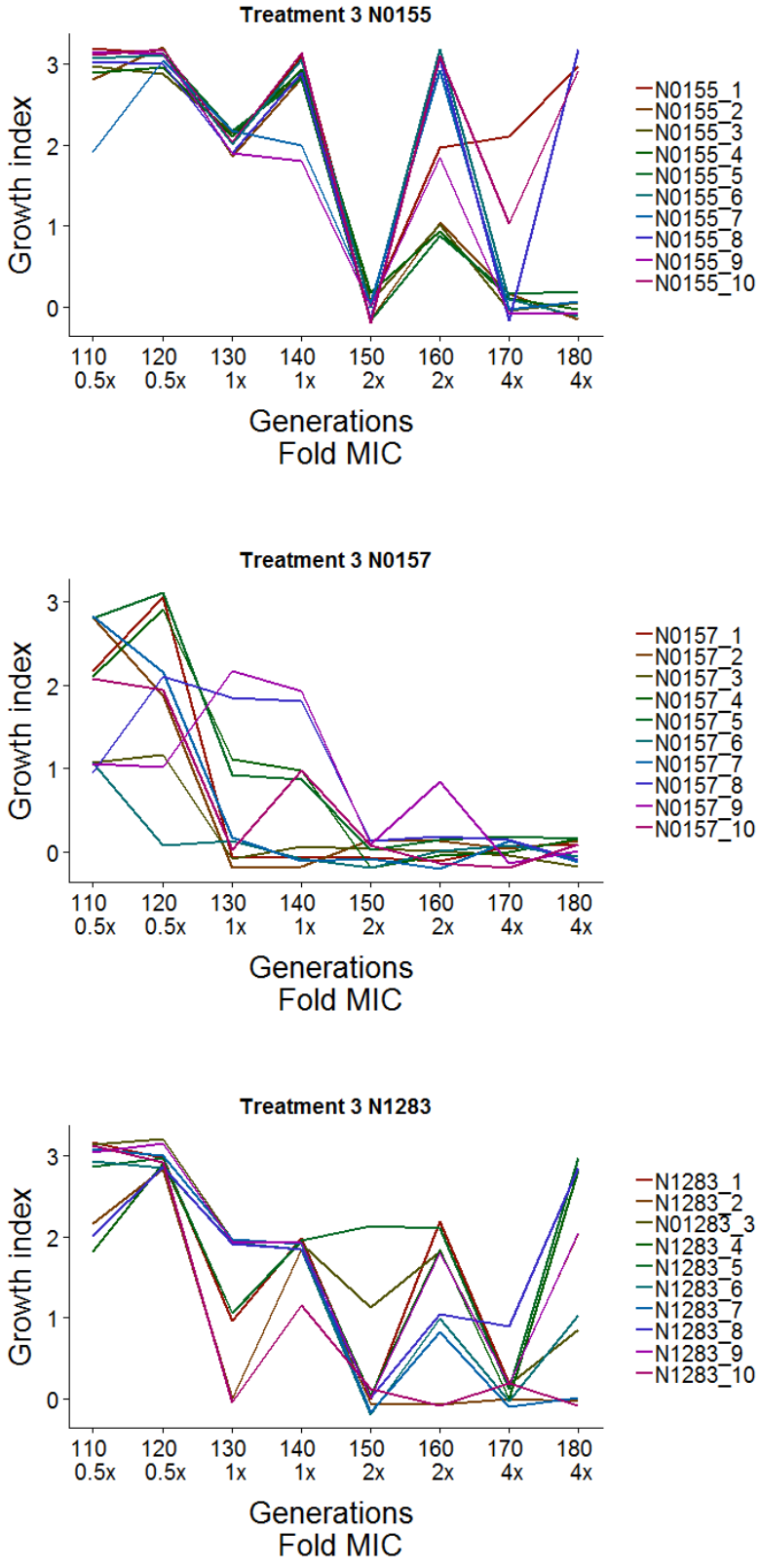

Figure 7.4.: **Growth indices of all cultures in treatment 3 (RIF + INH) starting from generation 110 (0.5 x MIC).**

differences in how the three strains react to increasing drug concentrations 1 (F= 79.4, p = 4.9 x 10^{-12}, repeated measurement ANOVA), treatment 2 (F = 38.7, p = 1.17 x 10^{-08}, repeated measurement ANOVA) and 3 (42.2 4.8 x 10^{-09}, repeated measurements ANOVA). The observed fluctuations were least pronounced in treatment 2, compared to treatments 1 and 3. Strains N0155 and N1283 do not seem to suffer from growth reduction in treatment 2, potentially indicating the early acquisition of mutations generating high-level INH resistance. Indeed, all sequenced endpoint cultures of treatment 2 harbour *katG* mutations. However, there are fluctuations in growth in strain N0157 in treatment 2, albeit on a high level (reduction from growth index 3 to 2). Mutations abrogating KatG function confer high-level INH resistance. The observed fluctuations in growth indices in N0157 in treatment 2 might be explained by *katG*, mutations which do not completely abrogate INH activation. Not all strains survived until the experiment was terminated in treatments 1 and 3 (table 7.2). Interestingly, not all strains from N1283 survived in treatment 1, but had the largest proportion of strains surviving in treatment 3. On the other hand, all N0157 replicates survived in treatment 1 but went extinct in treatment 3.

7.5. Discussion

After ≈180 generations of serial passage, we observe distinct patterns of growth between the different strains. For instance, RIF seems to exert a stronger influence on population density than INH, as demonstrated by the different growth patterns observed in treatments 1 and 3. This is not surprising, as the target size for *in vitro* INH resistance is larger compared to RIF (Bergval *et al.*, 2009); any mutation abrogating the activity of KatG will lead to INH resistance, whereas RIF resistance-conferring mutations are largely restricted to an 81 bp long region, termed the RIF resistance-determining region (Meftahi *et al.*, 2016).

The peculiar rebound pattern observed in all treatments containing RIF might be a sign of adaptation to higher drug concentrations not driven by resistance mutations. The growth dynamics in vented caps indicate that growth is comparable to agitated cultures when cell densities are low and the medium is sufficiently oxygenated for optimal growth. Changes in the endpoint cell numbers are therefore presumably due to inhibition action of the drugs in the early growth phases in the standing cultures. RIF degrades in 7H9 medium with a half-life of about a week (Yu *et al.*, 2011) but is not affected by the presence of INH. The inhibitory effect of RIF is therefore a function of time,

potentially giving strains the opportunity to survive either by upregulating drug efflux mechanisms, or by entering a non-replicative state until the RIF levels have degraded sufficiently (Brake *et al.*, 2018; Machado *et al.*, 2012). The observation of *katG* mutations in cultures that have never experienced INH is unexpected and prompted us to assess the possibility of experimental error including mislabelling vessels and tubes or swapping of experimental treatments. Mislabelling of the tubes used to prepare DNA for exploratory sequencing can be easily addressed by performing confirmatory PCRs on the replicates of treatment 1 carrying *katG* mutations. However, the presence of *rpoB* and *katG* double mutants in treatment 1 and differing *katG* mutations in the strains from treatment 2 (INH) make the swapping of labels between treatments 1 and 2 unlikely. The presence of double mutants in N0157 in treatment 1 also makes swapping labels between treatment 1 and 3 unlikely, as no cultures of N0157 survived the combination treatment. The appearance of double mutants might point towards swapping of experimental treatments, thus generating periodic selection potentially subsequently selecting for double resistant strains. However, swapping of experimental treatment would not be reconcilable with the observed growth indices, as these show repeating patterns. Furthermore, we do not observe *rpoB* mutations in the INH treatment, making it unlikely that the experimental treatments of treatment 1 and 2 were swapped.

If we exclude the possibility of experimental error, the presence of *katG* mutation in treatment 1 raises questions about their potential function in RIF resistance. Although we cannot be sure that all detected *katG* mutations in treatment 1 actually confer INH resistance, it is striking that 5/9 replicates in treatment 1 show *katG* mutations, whereas in the control experiment (no drug), only 1 strain acquired a *katG* mutation. The mechanism of INH resistance *in vitro* differs from that observed in clinico (Bergval *et al.*, 2009). INH resistance *in vitro* is conferred by any loss of function mutation in *katG*, whereas a functional *katG* is essential for pathogenicity of *M. tuberculosis in vivo* (Barnett *et al.*, 1953). The appearance of *katG* mutants in treatment 1 might indicate that the *katG* mutants have an altered susceptibility to RIF, presumably conferring the ability able to survive in the sub-MIC RIF environment. Unfortunately, there is only scarce quantitative drug susceptibility testing data assessing the RIF MIC of INH mono-resistant *katG* mutant strains. However, comparing RIF MIC data of 4 *katG* mutants with *katG* wild types presented in Chapter 4 indicates that the mean MIC of *katG* mutants is slightly higher (mean RIF MIC of *katG* mutants = 0.108 mg/L 7H10 agar dilution, mean RIF MIC *katG* wt = 0.95 mg/L, p = 0.051, Mann-Whitney test), although the difference only approaches significance and there are very few data points. The elevation of the RIF MIC in *katG*

mutants might have gone unnoticed, as the elevation of RIF MIC in *katG* mutants is subtle and might not cause elevations beyond the critical concentration.

The gene *katG* encodes a catalase/peroxidase involved in detoxification of reactive oxygen species. Losing the activity of this key detoxification enzyme might lead to the accumulation of toxic compounds and the subsequent induction of efflux mechanisms, potentially causing a slightly elevated RIF MIC. Indeed, efflux pumps have been demonstrated to be upregulated in MDR *M. tuberculosis* variants (Li *et al.*, 2015) and upon exposure of susceptible strains to the critical concentration of multiple drugs (Machado *et al.*, 2017). Furthermore, upregulation of efflux pumps seems to persist even after the acquisition of chromosomal mutations (Machado *et al.*, 2012). In treatment 3, the subtle increase in RIF MIC in *katG* mutants might aid the survival of INH mono-resistant strains in host tissues experiencing fluctuations in drug concentrations including sub-MIC RIF environments, especially in the light of current dosing regimens failing to generate sterilizing serum concentrations of RIF in certain patient groups (Chapter 3). Further evidence of the RIF MIC elevating effect of *katG* comes from the observation that in treatment 1, *rpoB* mutations conferring low-level RIF resistance (RpoB L430R, D435G), (Chapter 4, (Miotto *et al.*, 2018)) are only found in strains carrying *katG* mutations (3/9 strains). On the other hand in single mutants, *rpoB* mutations conferring high-level RIF resistance only RpoB S450L is found (2/9 strains). Given the frequent occurrence of *katG* mutants in treatment 1, it is not surprising that we find strains in treatment 3 without an *rpoB* mutation (2/6 strains). However, the detection of 2 *katG* mutants in N0157 in treatment 3 at generation 130 does not seem to fit the hypothesis that *katG* mutations increase the baseline RIF MIC, as all of the N0157 replicates in treatment 3 went extinct.

Interestingly, a recent study demonstrated that even small differences in RIF and INH MICs of otherwise susceptible strains are able to predict treatment failure and the evolution of drug resistant phenotypes in *M. tuberculosis* (Colangeli *et al.*, 2018). The observation that the distribution of *katG* mutations in clinico is skewed towards the S315T mutation (Chapter 5, Casali *et al.*, 2014; Merker *et al.*, 2018), rarely observed *in vitro*, might reflect the requirement of acquiring INH resistance while retaining catalase/peroxidase activity but might further offer a selective benefit by elevating the RIF MIC.

7.6. Future directions

First, we want to exclude experimental error by 1) confirming the fixed mutations identified by whole genome sequencing by Sanger sequencing from DNA directly isolated

frozen stocks from two consecutive freezer checkpoints (e.g. generation 160 & 180) and 2) sequencing the remainder of surviving endpoint cultures from generation 180.

7.6.1. Hypotheses

Assuming that experimental error is not the source of the reported observations, we want to address the following hypotheses:

1. *katG* mutations elevate the RIF MIC.
2. Elevated RIF MICs are the result of efflux pump upregulation
3. Efflux pump upregulation is a consequence of oxidative stress
4. *katG* mutations precede *rpoB* mutations in treatment 3 and treatment 1

7.6.2. Objectives

We want to test these hypotheses by addressing the following objectives:

1. Determine fine-scale RIF MICs (Colangeli *et al.*, 2018) for strains from treatment 1, 2 and 3 only carrying *katG* mutations. Furthermore, determine RIF MICs for strains which went through a single selection step for INH resistance, to exclude the possibility of other mutations acquired during serial passage, affecting RIF MICs
2. Determine fine-scale RIF MICs of *katG* mutants from treatment 1, 2 and 3 in presence and absence of efflux pump inhibitors (e.g. verapamil or reserpine)
3. Perform transcriptomic analyses of *katG* mutants and wild-type strains in presence / absence of sub-MIC RIF.
4. Measure fine-scale RIF MICs of *katG* mutant and wild-type strains in medium without additional catalase and/or medium supplemented with H_2O_2.
5. Sequence replicates from treatments 1 and 3 from a) generations predating the observed inhibitory effect of RIF (e.g. generation 80), b) generation (100 – growth inhibition becomes apparent)

7.6.3. Expected outcomes

If *katG* mutations elevate RIF MICs, we should be able to detect these with fine-scale MIC measurements. If these elevations are due to the induction of efflux pumps, the addition

of efflux pump inhibitors should reduce the RIF MIC again. Transcriptomic analyses will aid in detecting the putative cause of elevated MICs in *katG* mutants. If the elevation in RIF MICs is driven by oxidative-stress induced efflux mechanisms, the alteration in MIC might correlate with the level of oxidative stress, testable by varying the amount of H_2O_2 added to the culture medium. Finally, if *katG* mutations confer some resistance to RIF, we would expect them to evolve prior to RIF resistance mutations, as the target size for INH resistance *in vitro* is considerably larger (2.2 kb) than for RIF (95 % of RIF resistance mutations are within an 81 bp region of *rpoB*).

8. General discussion

The vast majority of TB cases are caused by drug-susceptible *M.tuberculosis* variants. However, certain parts of the world are heavily affected by drug-resistant variants, either due to the sheer number of incident TB cases in countries with large population sizes (e.g. China, India, The Democratic Republic of the Congo) or due to high proportions of cases caused by drug resistant variants among new cases (e.g. Georgia, Russia, South Africa) (World Health Organization, 2018a). There is an urgent need to better understand the ecology and evolution of drug resistance in *M. tuberculosis* to halt the spread and de novo evolution of drug resistant variants. This thesis examined multiple aspects of the biology of drug resistance in TB relevant for detecting drug resistant variants and predicting quantitative levels of drug resistance from the genotype (Chapter 4), understanding how drug resistant variants spread (Chapter 5), understanding the factors influencing the transmission of resistant variants (Chapter 6) and examining conditions which favour the evolution of multidrug-resistant variants (Chapter 7).

8.1. Remaining challenges for implementing WGS in routine drug susceptibility testing in TB

In Chapter 4, we investigated the utility of whole genome sequencing for the detection of drug-resistant *M.tuberculosis* variants, as well as for predicting quantitative levels of drug resistance. The early and timely detection of drug resistant variants is a prerequisite for ensuring adequate treatment of patients and limiting the spread of drug resistant variants, benefiting individual patients and populations at large. Despite the fact that we were able to show that WGS is a powerful tool for detecting drug resistance-and partially predicting quantitative levels of drug resistance, there are still some challenges ahead for implementing the technology in routine diagnostic use.

8.1.1. Standardization of WGS analysis for DST

The reproducibility of WGS-based DST may in part depend on the methods used to identify mutations, as well as the sequencing technology used to generate the sequencing data (Phelan *et al.*, 2016). These observations call for standardized approaches or at least establishment of best practices to sequence and identify mutations for WGS-based DST in the clinics. Several standardized pipelines have been published in recent years. A systematic review correlating phenotypic DST data with mutations identified by the different pipelines would help to establish optimal procedures to obtain the optimal sensitivity and specificity for WGS-based DST in *M.tuberculosis*. However, the diagnostic power of WGS-based DST is not only dependent on reproducible analysis strategies, but also on the ability to separate phylogenetic variation from mutations causing drug resistance. To avoid false positive results, there is a need for a comprehensive database of phylogenetic markers not involved in conferring resistance. This requires large-scale sequencing efforts combined with phenotypic DST. Such efforts are well under way to resolve many issues related to the presence of phylogenetic markers in drug resistance genes (Starks *et al.*, 2015; The CRyPTIC Consortium and the 100000 Genomes Project, 2018).

8.1.2. Reducing diagnostic delay

Reducing the diagnostic delay is crucial to provide effective treatment to patients and subsequently reduce transmission of TB. Sequencing-based DST is a powerful tool to detect drug-resistance conferring mutations and, compared to phenotypic DST methods, comes with a much shorter turnaround time. However, it still requires a culture step to generate sufficient DNA for sequencing. This culture step is time- consuming and skews the sample towards cells which were able to grow in the medium, reducing sample diversity and potentially losing drug resistant variants, which would lead to false-negative results. Sequencing directly from patient sputum eliminates the culture step and retains a maximum in diversity, as the cells present in the sputum do not have to be viable/culturable and (non-dormant) (Chengalroyen *et al.*, 2016). Considerable advances on sequencing directly from patient sputa have been made recently, which constitute important steps in eliminating the need for a culture step (Brown *et al.*, 2015).

8.1.3. Limitations of WGS for DST in *M. tuberculosis*

We observed considerable variability in minimal inhibitory concentrations conferred by identical drug resistance mutations for a number of drugs. This variability might be ex-

plained by epistatic interactions between the resistance mutations and the genetic background and would in turn be amenable to study using WGS. However, it is becoming increasingly apparent that non-mutational mechanisms may influence minimal inhibitory concentrations (MIC). Specifically, the plastic expression of efflux mechanisms is involved in conferring clinically relevant increases in MIC (Mazando *et al.*, 2017). Differences in efflux pump expression are not detectable with WGS. However, drug-resistant strains with upregulated efflux pumps usually also harbour bona fide drug resistance mutations, indicating that the expression of efflux pumps might act as a stepping stone in evolving high-level drug resistance. More in-depth analyses are required to establish or refute a connection between efflux pumps and elevated MICs in absence of genetic resistance mechanisms.

8.2. Drug combinations and the evolution of drug resistance

TB infections are treated with a combination of drugs, requiring us to view the evolution of resistance against single drugs in the in the context of combination treatment. In Chapter 7, we report the peculiar observation that mutations in the gene *katG*, which are typically associated with INH resistance, seem to confer low-level RIF resistance. It seems likely that the elevated RIF MICs in *katG* mutants are driven by drug efflux mechanisms. If this observation can be confirmed, it will prompt us to investigate how resistance mutations in general influence MICs of drugs via the expression of efflux mechanisms, potentially generating an environment permissive for the evolution of strains resistant to multiple drugs. Furthermore, it highlights the necessity to study how sub-MIC environments are generated within the body and how one can prevent them. In the case of RIF, the solution would be fairly easy, as one could increase the dose of RIF administered. Clinical trials have already demonstrated that increased doses of RIF are safe and have better treatment outcomes (Boeree *et al.*, 2017). Given the mounting evidence that the plastic expression of efflux pumps has clinical consequences, it might be wise to study whether efflux pump inhibitors might slow the evolution of resistant strains (Pule *et al.*, 2016). In the light of novel antituberculous drugs bedaquiline, pretomanid and delamanid (Zumla *et al.*, 2014) being used to treat MDR-TB, as well as being part of novel treatment regimens, we need to remain exceptionally vigilant to ensure that these drugs do not lose their utility early on due to the evolution of resistance. For this, we need to understand

the resistance mechanisms in detail, especially how sub-MIC concentrations affect the evolution of resistance against these novel drugs.

8.3. From genome sequences to public health interventions

Going beyond predicting drug resistance profiles in *M.tuberculosis*, WGS is exceptionally useful to infer transmission chains (Walker et al. 2013; Stucki et al. 2016a). However, phylogenetic models are inadequate to describe relationships between highly similar strains frequently found in pathogen transmission studies. Strains often do not display any sequence differences, violating the key assumption of shared ancestry in phylogenies. Furthermore, methods solely relying on genetic distance are unable to infer the directionality of transmission events. New methods (Klinkenberg et al. 2017) allow for the simultaneous inference of phylogenetic and transmission trees, by incorporating models on the time course of the infection with isolation dates. In Chapter 5, we were able to demonstrate the utility of WGS to infer transmission using a nation-wide and population-based collection of MDR-TB strains from Georgia. We were able to demonstrate that prisons stand at the centre of the MDR-TB epidemic in Georgia, as well as quantify the overall effect of prisons on case numbers of MDR-TB outside of prisons. There is reason to believe that the situation in Georgia can serve as a proxy for many countries of the former Soviet Union (Droznin *et al.*, 2017). We uncovered one of the largest transmission clusters of pre-extensively drug resistant TB ever reported, reflecting the difficulties prisoners have in accessing adequate medical treatment (Department of State of the United States of America, 2012). It is puzzling how such a large transmission cluster could have gone unnoticed. In light of the frequently observed transmission in Georgian prisons, there is an urgent need for public health interventions. In addition to providing adequate treatment within prisons (World Health Organization, 2015), follow-up of former prison inmates diagnosed with TB is an urgent task to tackle. Former prison inmates frequently have low socio-economic status and lack resources to access medical services after being released from prison and should be actively followed up. Concerted public health interventions in settings frequently found in the former Soviet Union can be highly effective, as demonstrated by the rigorous implementation of the WHO End-TB strategy in Georgia's neighbour country Azerbaijan. TB incidence in Azerbaijani prisons dropped by a factor of 2.5 in the past 15 years (Gurbanova *et al.*, 2016). Public health interventions require urgent governmental action and funding, especially since the TB control program in Geor-

gia was funded in part by the Global Fund, which is phasing out its financial support in the coming years (World Health Organization, 2015).

8.3.1. Whole genome sequencing for real-time monitoring of the TB epidemic

The cost of sequencing has decreased dramatically over the past years (https://www.genome.gov/sequencingcostsdata/), creating a number of intriguing opportunities. Together with the advances in laboratory automatisation, the decreased sequencing costs will make it feasible to routinely sequence all strains isolated from patients. Platforms like Nextstrain (https://nextstrain.org/) will be exceptionally useful, as they allow putting sequences to be put directly into the context of all other submitted sequence data. This will allow for the direct detection of transmission clusters in real time and aid to in direct active case finding campaigns, greatly increasing the ability of public health systems to interrupt transmission and reduce the TB burden. Occurrences of transmission clusters of the size described in Chapter 5 would be unlikely to occur. Data like these will offer unprecedented insight into pathogen biology and help to monitor public health system interventions. Not only will we be able to monitor public health interventions, we will be able to measure how human interventions influence the evolutionary trajectories of pathogens.

8.4. Population-based samples and the study of adaptation

Even if routine sequencing of all patient isolates will not be feasible in the short term, sampling efforts in Georgia continue. The National Centre of Tuberculosis and Lung Diseases in Tbilisi is collecting and storing all patient isolates (ca. 3-4000 strains per year) from 2013 onwards, enabling the study of the whole population when sequencing costs have decreased sufficiently. Such a sample set will allow researchers to study the compensation of drug resistance-related fitness costs with the best possible resolution using network analysis-based approaches such as those outlined in Chapter 6. The advent of long-read sequencing technologies will make the reliable and in-depth study of structural genome variation feasible. In the absence of detectable horizontal gene transfer, structural variation is believed to be an important source of variation in *M.tuberculosis* (Gagneux *et al.*, 2006c). Furthermore, in Chapter 5, we had to exclude a number of strains due to an

excess of unfixed positions, indicating either cross-contamination of cultures in the lab or mixed infections (Cohen *et al.*, 2012). Given the high transmission rates observed within prisons, infections with multiple strains appear possible. Third-generation sequencers will offer the possibility to reliably reconstruct haplotypes. Combined with high-density population-based samples we will be able to study how mixed infections shape the TB epidemic.

8.5. Conclusions

Despite the overall declining incidence of TB (World Health Organization, 2018a), the epidemic is far from being under control. Furthermore, the extensive prevalence of highly drug resistant strains, especially in countries of the former Soviet Union, calls for extensive and tightly monitored public health interventions. Whole-genome sequencing can provide the means for generating accurate and timely antibiograms, as well as understanding transmission of highly resistant strains. Furthermore, population-based strain collections offer the possibility to monitor the impact of public health interventions. However, we still lack a comprehensive understanding of which genetic determinants make drug-resistant strains transmit efficiently. This understanding will help us to develop methods to detect strains with the capability to efficiently generate secondary cases. Finally, understanding the effects of sub-MIC drug combinations on the evolution of drug resistance, with a particular focus on the influence of drug efflux mechanisms, will help us to develop interventions to slow or prevent the evolution of highly resistant *M.tuberculosis* strains.

9. Bibliography

Abate, G, Hoffner, SE (1997) Synergistic antimycobacterial activity between ethambutol and the beta-lactam drug cefepime. *Diagnostic Microbiology and Infectious Disease*, **28**(3): 119–122.

Abraham, E, Chain, E, Fletcher, C, Gardner, A, Heatley, N, Jennings, M, Florey, H (1941) FURTHER OBSERVATIONS ON PENICILLIN. *The Lancet*, **238**(6155): 177–189.

Achtman, M (2008) Evolution, Population Structure, and Phylogeography of Genetically Monomorphic Bacterial Pathogens. *Annual Review of Microbiology*, **62**(1): 53–70.

Adams, KN, Takaki, K, Connolly, LE, Wiedenhoft, H, Winglee, K, Humbert, O, Edelstein, PH, Cosma, CL, Ramakrishnan, L (2011) Drug Tolerance in Replicating Mycobacteria Mediated by a Macrophage-Induced Efflux Mechanism. *Cell*, **145**(1): 39–53.

Aerts, A, Habouzit, M, Mschiladze, L, Malakmadze, N, Sadradze, N, Menteshashvili, O, Portaels, F, Sudre, P (2000) Pulmonary tuberculosis in prisons of the ex-USSR state Georgia: results of a nation-wide prevalence survey among sentenced inmates. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease*, **4**(12): 1104–10.

Ahmad, N *et al.* (2018) Treatment correlates of successful outcomes in pulmonary multidrug-resistant tuberculosis: an individual patient data meta-analysis. *The Lancet*, **392**(10150): 821–834.

Alangaden, GJ, Kreiswirth, BN, Aouad, A, Khetarpal, M, Igno, FR, Moghazeh, SL, Manavathu, EK, Lerner, SA (1998) Mechanism of resistance to amikacin and kanamycin in Mycobacterium tuberculosis. *Antimicrobial agents and chemotherapy*, **42**(5): 1295–7.

Andersson, DI, Levin, BR (1999) The biological cost of antibiotic resistance. *Current Opinion in Microbiology*, **2**(5): 489–493.

Andersson, DI, Hughes, D (2010) Antibiotic resistance and its cost: is it possible to reverse resistance? *Nature Reviews Microbiology*, **8**(4): 260–71.

Andersson, DI, Hughes, D (2014) Microbiological effects of sublethal levels of antibiotics. *Nature reviews. Microbiology*, **12**(7): 465–78.

Andre, E, Goeminne, L, Cabibbe, A, Beckert, P, Kabamba Mukadi, B, Mathys, V, Gagneux, S, Niemann, S, Van Ingen, J, Cambau, E (2017) Consensus numbering system for the rifampicin resistance-associated rpoB gene mutations in pathogenic mycobacteria. *Clinical Microbiology and Infection*, **23**(3): 167–172.

Ängeby, K, Juréen, P, Kahlmeter, G, Hoffner, S, Schön, T (2012) Challenging a dogma: antimicrobial susceptibility testing breakpoints for Mycobacterium tuberculosis. *Bulletin of the World Health Organization*, **90**(9): 693–698.

Ankomah, P, Levin, BR (2012) Two-drug antimicrobial chemotherapy: a mathematical model and experiments with Mycobacterium marinum. *PLoS pathogens*, **8**(1): e1002487.

Anthony Malinga, L, Stoltz, A (2016) Efflux Pump Mediated Second-Line Tuberculosis Drug Resistance. *Mycobacterial Diseases*, **6**(3):

Atun, R, Weil, DE, Eang, MT, Mwakyusa, D (2010) Health-system strengthening and tuberculosis control. *The Lancet*, **375**(9732): 2169–2178. eprint: ISSN0042-9686.

Banerjee, A, Dubnau, E, Quemard, A, Balasubramanian, V, Um, K, Wilson, T, Collins, D, Lisle, G de, Jacobs, W (1994) inhA, a gene encoding a target for isoniazid and ethionamide in Mycobacterium tuberculosis. *Science*, **263**(5144): 227–230.

Barnett, M, Busby, SR, Mitchison, DA (1953) Tubercle bacilli resistant to isoniazid: virulence and response to treatment with isoniazid in guinea-pigs and mice. *British journal of experimental pathology*, **34**(5): 568–81.

Baussano, I, Williams, BG, Nunn, P, Beggiato, M, Fedeli, U, Scano, F (2010) Tuberculosis incidence in prisons: A systematic review. *PLoS Medicine*, **7**(12): 1–10.

Beckert, P, Hillemann, D, Kohl, TA, Kalinowski, J, Richter, E, Niemann, S, Feuerriegel, S (2012) rplC T460C Identified as a Dominant Mutation in Linezolid-Resistant Mycobacterium tuberculosis Strains. *Antimicrobial Agents and Chemotherapy*, **56**(5): 2743–2745.

Behr, MA, Edelstein, PH, Ramakrishnan, L (2018) Revisiting the timetable of tuberculosis. *BMJ*, **2738**(August): k2738.

Bergval, IL, Schuitema, ARJ, Klatser, PR, Anthony, RM (2009) Resistant mutants of Mycobacterium tuberculosis selected in vitro do not reflect the in vivo mechanism of isoniazid resistance. *The Journal of antimicrobial chemotherapy*, **64**(3): 515–23.

Bergval, I, Kwok, B, Schuitema, A, Kremer, K, Soolingen, D van, Klatser, P, Anthony, R (2012) Pre-existing isoniazid resistance, but not the genotype of Mycobacterium tuberculosis drives rifampicin resistance codon preference in vitro. *PloS one*, **7**(1): e29108.

Bernard, C, Aubry, A, Chauffour, A, Brossier, F, Robert, J, Veziris, N (2016) In vivo Mycobacterium tuberculosis fluoroquinolone resistance emergence: a complex phenomenon

poorly detected by current diagnostic tests. *Journal of Antimicrobial Chemotherapy*, **71**(12): 3465–3472.

Berrada, ZL, Lin, SYG, Rodwell, TC, Nguyen, D, Schecter, GF, Pham, L, Janda, JM, Elmaraachli, W, Catanzaro, A, Desmond, E (2016) Rifabutin and rifampin resistance levels and associated rpoB mutations in clinical isolates of Mycobacterium tuberculosis complex. *Diagnostic Microbiology and Infectious Disease*, **85**(2): 177–181.

Billington, OJ, McHugh, TD, Gillespie, SH (1999) Physiological cost of rifampin resistance induced in vitro in Mycobacterium tuberculosis. *Antimicrobial agents and chemotherapy*, **43**(8): 1866–9.

Björkman, J, Nagaev, I, Berg, OG, Hughes, D, Andersson, DI (2000) Effects of environment on compensatory mutations to ameliorate costs of antibiotic resistance. *Science (New York, N.Y.)*, **287**(5457): 1479–82.

Bloemberg, GV, Keller, PM, Stucki, D, Trauner, A, Borrell, S, Latshang, T, Coscolla, M, Rothe, T, Hömke, R, Ritter, C, Feldmann, J, Schulthess, B, Gagneux, S, Böttger, EC (2015) Acquired Resistance to Bedaquiline and Delamanid in Therapy for Tuberculosis. *New England Journal of Medicine*, **373**(20): 1986–1988.

Blower, SM, Chou, T (2004) Modeling the emergence of the 'hot zones': tuberculosis and the amplification dynamics of drug resistance. *Nature medicine*, **10**(10): 1111–6.

Boeree, MJ, Heinrich, N, Aarnoutse, R, Diacon, AH, Dawson, R, Rehal, S, Kibiki, GS, Churchyard, G, Sanne, I, Ntinginya, NE, Minja, LT, Hunt, RD, Charalambous, S, Hanekom, M, Semvua, HH, Mpagama, SG, Manyama, C, Mtafya, B, Reither, K, Wallis, RS, Venter, A, Narunsky, K, Mekota, A, Henne, S, Colbers, A, Balen, GP van, Gillespie, SH, Phillips, PPJ, Hoelscher, M (2017) High-dose rifampicin, moxifloxacin, and SQ109 for treating tuberculosis: a multi-arm, multi-stage randomised controlled trial. *The Lancet Infectious Diseases*, **17**(1): 39–49.

Bolotin, E, Hershberg, R (2015) Gene Loss Dominates As a Source of Genetic Variation within Clonal Pathogenic Bacterial Species. *Genome biology and evolution*, **7**(8): 2173–87.

Borrell, S, Gagneux, S (2011) Strain diversity, epistasis and the evolution of drug resistance in Mycobacterium tuberculosis. *Clinical Microbiology and Infection*, **17**(6): 815–820.

Borrell, S, Teo, Y, Giardina, F, Streicher, EM, Klopper, M, Feldmann, J, Müller, B, Victor, TC, Gagneux, S (2013) Epistasis between antibiotic resistance mutations drives the evolution of extensively drug-resistant tuberculosis. *Evolution, Medicine, and Public Health*, **2013**(1): 65–74.

Bosne-David, S (2000) Intrinsic resistance of Mycobacterium tuberculosis to clarithromycin is effectively reversed by subinhibitory concentrations of cell wall inhibitors. *Journal of Antimicrobial Chemotherapy*, **46**(3): 391–395.

Böttger, E (2011) The ins and outs of Mycobacterium tuberculosis drug susceptibility testing. *Clinical Microbiology and Infection*, **17**(8): 1128–1134.

Böttger, EC, Springer, B (2008) Tuberculosis: drug resistance, fitness, and strategies for global control. *European Journal of Pediatrics*, **167**(2): 141–148.

Böttger, EC, Springer, B, Pletschette, M, Sander, P (1998) Fitness of antibiotic-resistant microorganisms and compensatory mutations. *Nature Medicine*, **4**(12): 1343–1344.

Brake, LH te, Knegt, GJ de, Steenwinkel, JE de, Dam, TJ van, Burger, DM, Russel, FG, Crevel, R van, Koenderink, JB, Aarnoutse, RE (2018) The Role of Efflux Pumps in Tuberculosis Treatment and Their Promise as a Target in Drug Development: Unraveling the Black Box. *Annual Review of Pharmacology and Toxicology*, **58**(1): annurev–pharmtox–010617–052438.

Brandis, G, Wrande, M, Liljas, L, Hughes, D (2012) Fitness-compensatory mutations in rifampicin-resistant RNA polymerase. *Molecular Microbiology*, **85**(1): 142–151.

Brennan, PJ, Nikaido, H (1995) The Envelope of Mycobacteria. *Annual Review of Biochemistry*, **64**(1): 29–63.

Brites, D, Gagneux, S (2012) Old and new selective pressures on Mycobacterium tuberculosis. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases*, **12**(4): 678–85.

Brites, D, Gagneux, S (2015) Co-evolution of Mycobacterium tuberculosis and Homo sapiens. *Immunological Reviews*, **264**(1): 6–24.

British Thoracic and Tuberculosis Association (1975) Short-course chemotherapy in pulmonary tuberculosis. *The Lancet*, **305**(7899): 119–124.

Brown, AC, Bryant, JM, Einer-Jensen, K, Holdstock, J, Houniet, DT, Chan, JZ, Depledge, DP, Nikolayevskyy, V, Broda, A, Stone, MJ, Christiansen, MT, Williams, R, McAndrew, MB, Tutill, H, Brown, J, Melzer, M, Rosmarin, C, McHugh, TD, Shorten, RJ, Drobniewski, F, Speight, G, Breuer, J (2015) Rapid whole-genome sequencing of mycobacterium tuberculosis isolates directly from clinical samples. *Journal of Clinical Microbiology*, **53**(7): 2230–2237.

Buriánková, K, Doucet-Populaire, F, Dorson, O, Gondran, A, Ghnassia, JC, Weiser, J, Pernodet, JL (2004) Molecular basis of intrinsic macrolide resistance in the Mycobacterium tuberculosis complex. *Antimicrobial agents and chemotherapy*, **48**(1): 143–50.

Calver, AD, Falmer, AA, Murray, M, Strauss, OJ, Streicher, EM, Hanekom, M, Liversage, T, Masibi, M, Helden, PDV, Warren, RM, Victor, TC (2010) Emergence of Increased

Resistance and Extensively Drug-Resistant Tuberculosis Despite Treatment Adherence , South Africa. *Emerging Infectious Diseases*, **16**(2): 264–271.

Caminero, Ja (2008) Likelihood of generating MDR-TB and XDR-TB under adequate National Tuberculosis Control Programme implementation. *The international journal of tuberculosis and lung disease*, **12**(8): 869–77.

Campbell, EA, Korzheva, N, Mustaev, A, Murakami, K, Nair, S, Goldfarb, A, Darst, SA (2001) Structural mechanism for rifampicin inhibition of bacterial rna polymerase. *Cell*, **104**(6): 901–12.

Casali, N, Nikolayevskyy, V, Balabanova, Y, Harris, SR, Ignatyeva, O, Kontsevaya, I, Corander, J, Bryant, J, Parkhill, J, Nejentsev, S, Horstmann, RD, Brown, T, Drobniewski, F (2014) Evolution and transmission of drug-resistant tuberculosis in a Russian population. *Nature genetics*, **46**(3): 279–86.

Casali, N, Nikolayevskyy, V, Balabanova, Y, Ignatyeva, O, Kontsevaya, I, Harris, SR, Bentley, SD, Parkhill, J, Nejentsev, S, Hoffner, SE, Horstmann, RD, Brown, T, Drobniewski, F (2012) Microevolution of extensively drug-resistant tuberculosis in Russia. *Genome Research*, **22**(4): 735–745.

Chambers, HF, Moreau, D, Yajko, D, Miick, C, Wagner, C, Hackbarth, C, Kocagoz, S, Rosenberg, E, Hadley, WK, Nikaido, H (1995) Can penicillins and other beta-lactam antibiotics be used to treat tuberculosis? *Antimicrobial Agents and Chemotherapy*, **39**(12): 2620–2624.

Chen, PE, Shapiro, BJ (2015) The advent of genome-wide association studies for bacteria. *Current Opinion in Microbiology*, **25** 17–24.

Chengalroyen, MD, Beukes, GM, Gordhan, BG, Streicher, EM, Churchyard, G, Hafner, R, Warren, R, Otwombe, K, Martinson, N, Kana, BD (2016) Detection and Quantification of Differentially Culturable Tubercle Bacteria in Sputum from Patients with Tuberculosis. *American Journal of Respiratory and Critical Care Medicine*, **194**(12): 1532–1540.

Chou, HH, Chiu, HC, Delaney, NF, Segre, D, Marx, CJ (2011) Diminishing Returns Epistasis Among Beneficial Mutations Decelerates Adaptation. *Science*, **332**(6034): 1190–1192.

Cohen, KA, Abeel, T, Manson McGuire, A, Desjardins, CA, Munsamy, V, Shea, TP, Walker, BJ, Bantubani, N, Almeida, DV, Alvarado, L, Chapman, SB, Mvelase, NR, Duffy, EY, Fitzgerald, MG, Govender, P, Gujja, S, Hamilton, S, Howarth, C, Larimer, JD, Maharaj, K, Pearson, MD, Priest, ME, Zeng, Q, Padayatchi, N, Grosset, J, Young, SK, Wortman, J, Mlisana, KP, O'Donnell, MR, Birren, BW, Bishai, WR, Pym, AS, Earl, AM (2015) Evolution of Extensively Drug-Resistant Tuberculosis over Four

Decades: Whole Genome Sequencing and Dating Analysis of Mycobacterium tuberculosis Isolates from KwaZulu-Natal. *PLOS Medicine*, **12**(9): ed. by JZ Metcalfe, e1001880.

Cohen, KA, El-Hay, T, Wyres, KL, Weissbrod, O, Munsamy, V, Yanover, C, Aharonov, R, Shaham, O, Conway, TC, Goldschmidt, Y, Bishai, WR, Pym, AS (2016) Paradoxical Hypersusceptibility of Drug-resistant Mycobacterium tuberculosis to β-lactam Antibiotics. *EBioMedicine*, **9** 170–179.

Cohen, T, Helden, PD van, Wilson, D, Colijn, C, McLaughlin, MM, Abubakar, I, Warren, RM (2012) Mixed-strain mycobacterium tuberculosis infections and the implications for tuberculosis treatment and control. *Clinical microbiology reviews*, **25**(4): 708–19.

Cohen, T, Murray, M (2004) Modeling epidemics of multidrug-resistant M. tuberculosis of heterogeneous fitness. *Nature medicine*, **10**(10): 1117–21.

Colangeli, R, Jedrey, H, Kim, S, Connell, R, Ma, S, Chippada Venkata, UD, Chakravorty, S, Gupta, A, Sizemore, EE, Diem, L, Sherman, DR, Okwera, A, Dietze, R, Boom, WH, Johnson, JL, Mac Kenzie, WR, Alland, D (2018) Bacterial Factors That Predict Relapse after Tuberculosis Therapy. *New England Journal of Medicine*, **379**(9): 823–833.

Cole, ST, Brosch, R, Parkhill, J, Garnier, T, Churcher, C, Harris, D, Gordon, SV, Eiglmeier, K, Gas, S, Barry, CE, Tekaia, F, Badcock, K, Basham, D, Brown, D, Chillingworth, T, Connor, R, Davies, R, Devlin, K, Feltwell, T, Gentles, S, Hamlin, N, Holroyd, S, Hornsby, T, Jagels, K, Krogh, a, McLean, J, Moule, S, Murphy, L, Oliver, K, Osborne, J, Quail, Ma, Rajandream, Ma, Rogers, J, Rutter, S, Seeger, K, Skelton, J, Squares, R, Squares, S, Sulston, JE, Taylor, K, Whitehead, S, Barrell, BG (1998) Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature*, **393**(6685): 537–544.

Coleman, MT, Chen, RY, Lee, M, Lin, PL, Dodd, LE, Maiello, P, Via, LE, Kim, Y, Marriner, G, Dartois, V, Scanga, C, Janssen, C, Wang, J, Klein, E, Cho, SN, Barry, CE, Flynn, JL (2014) PET/CT imaging reveals a therapeutic response to oxazolidinones in macaques and humans with tuberculosis. *Science translational medicine*, **6**(265): 265ra167.

Colijn, C, Cohen, T, Ganesh, A, Murray, M (2011) Spontaneous emergence of multiple drug resistance in tuberculosis before and during therapy. *PloS one*, **6**(3): e18327.

Coll, F, McNerney, R, Preston, MD, Guerra-Assunção, JA, Warry, A, Hill-Cawthorne, G, Mallard, K, Nair, M, Miranda, A, Alves, A, Perdigão, J, Viveiros, M, Portugal, I, Hasan, Z, Hasan, R, Glynn, JR, Martin, N, Pain, A, Clark, TG (2015) Rapid determination of anti-tuberculosis drug resistance from whole-genome sequences. *Genome Medicine*, **7**(1): 51.

Colman, RE, Anderson, J, Lemmer, D, Lehmkuhl, E, Georghiou, SB, Heaton, H, Wiggins, K, Gillece, JD, Schupp, JM, Catanzaro, DG, Crudu, V, Cohen, T, Rodwell, TC, Engelthaler, DM (2016) Rapid Drug Susceptibility Testing of Drug-Resistant Mycobacterium tuberculosis Isolates Directly from Clinical Samples by Use of Amplicon Sequencing: a Proof-of-Concept Study. *Journal of Clinical Microbiology*, **54**(8): ed. by GA Land, 2058–2067.

Comas, I, Borrell, S, Roetzer, A, Rose, G, Malla, B, Kato-Maeda, M, Galagan, J, Niemann, S, Gagneux, S (2012) Whole-genome sequencing of rifampicin-resistant Mycobacterium tuberculosis strains identifies compensatory mutations in RNA polymerase genes. *Nature Genetics*, **44**(1): 106–110.

Coscolla, M, Gagneux, S (2014) Consequences of genomic diversity in Mycobacterium tuberculosis. *Seminars in Immunology*, **26**(6): 431–444. eprint: NIHMS150003.

Coscolla, M, Gagneux, S (2010) Does M. tuberculosis genomic diversity explain disease diversity? *Drug discovery today. Disease mechanisms*, **7**(1): e43–e59.

Crofton, J, Mitchison, D (1948) Streptomycin Resistance in Pulmonary Tuberculosis. *British Medical Journal*, **2(4588)** 1009–1015.

Danilchanka, O, Pires, D, Anes, E, Niederweis, M (2015) The Mycobacterium tuberculosis Outer Membrane Channel Protein CpnT Confers Susceptibility to Toxic Molecules. *Antimicrobial Agents and Chemotherapy*, **59**(4): 2328–2336.

Danilchanka, O, Sun, J, Pavlenok, M, Maueroder, C, Speer, A, Siroy, A, Marrero, J, Trujillo, C, Mayhew, DL, Doornbos, KS, Munoz, LE, Herrmann, M, Ehrt, S, Berens, C, Niederweis, M (2014) An outer membrane channel protein of Mycobacterium tuberculosis with exotoxin activity. *Proceedings of the National Academy of Sciences*, **111**(18): 6750–6755.

Dartois, V (2014) The path of anti-tuberculosis drugs: from blood to lesions to mycobacterial cells. *Nature Reviews Microbiology*, **12**(3): 159–167. eprint: NIHMS150003.

De Lorenzo, S, Alffenaar, JW, Sotgiu, G, Centis, R, D'Ambrosio, L, Tiberi, S, Bolhuis, MS, Van Altena, R, Viggiani, P, Piana, A, Spanevello, A, Migliori, GB (2013) Efficacy and safety of meropenem-clavulanate added to linezolid-containing regimens in the treatment of MDR-/XDR-TB. *European Respiratory Journal*, **41**(6): 1386–1392.

Deggim-Messmer, V, Bloemberg, GV, Ritter, C, Voit, A, Hömke, R, Keller, PM, Böttger, EC (2016) Diagnostic Molecular Mycobacteriology in Regions With Low Tuberculosis Endemicity: Combining Real-time PCR Assays for Detection of Multiple Mycobacterial Pathogens With Line Probe Assays for Identification of Resistance Mutations. *EBioMedicine*, **9** 228–237.

Delignette-Muller, ML, Dutang, C (2015) fitdistrplus: an R package for fitting distributions. *Journal of Statistical Software*, **64**(4): 1–34.

Department of State of the United States of America (2012) *GEORGIA 2012 Human Rights Report*. Tech. rep. Washington DC: United States Department of State, pp. 1–61.

Desjardins, CA, Cohen, KA, Munsamy, V, Abeel, T, Maharaj, K, Walker, BJ, Shea, TP, Almeida, DV, Manson, AL, Salazar, A, Padayatchi, N, O'Donnell, MR, Mlisana, KP, Wortman, J, Birren, BW, Grosset, J, Earl, AM, Pym, AS (2016a) Genomic and functional analyses of Mycobacterium tuberculosis strains implicate ald in D-cycloserine resistance. *Nature Genetics*, **48**(5): 544–551.

Desjardins, CA, Cohen, KA, Munsamy, V, Abeel, T, Maharaj, K, Walker, BJ, Shea, TP, Almeida, DV, Manson, AL, Salazar, A, Padayatchi, N, O'Donnell, MR, Mlisana, KP, Wortman, J, Birren, BW, Grosset, J, Earl, AM, Pym, AS (2016b) Genomic and functional analyses of Mycobacterium tuberculosis strains implicate ald in D-cycloserine resistance. *Nature Genetics*, **48**(5): 544–551.

Dheda, K, Barry, CE, Maartens, G (2016) Tuberculosis. *The Lancet*, **387**(10024): 1211–1226.

Dickinson, JM, Aber, VR, Mitchison, DA (1977) Bactericidal activity of streptomycin, isoniazid, rifampin, ethambutol, and pyrazinamide alone and in combination against Mycobacterium Tuberculosis. *The American review of respiratory disease*, **116**(4): 627–635.

Domínguez, J, Boettger, EC, Cirillo, D, Cobelens, F, Eisenach, KD, Gagneux, S, Hillemann, D, Horsburgh, R, Molina-Moya, B, Niemann, S, Tortoli, E, Whitelaw, A, Lange, C, TBNET, networks, RT (2016) Clinical implications of molecular drug resistance testing for Mycobacterium tuberculosis: a TBNET/RESIST-TB consensus statement. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease*, **20**(1): 24–42.

Doorn, H van, Haas, P de, Kremer, K, Vandenbroucke-Grauls, C, Borgdorff, M, Soolingen, D van (2006) Public health impact of isoniazid-resistant Mycobacterium tuberculosis strains with a mutation at amino-acid position 315 of katG: a decade of experience in The Netherlands. *Clinical Microbiology and Infection*, **12**(8): 769–775.

Droznin, M, Johnson, A, Johnson, AM (2017) Multidrug resistant tuberculosis in prisons located in former Soviet countries: A systematic review. *PLOS ONE*, **12**(3): ed. by R Manganelli, e0174373.

Dye, C, Espinal, MA (2001) Will tuberculosis become resistant to all antibiotics? *Proceedings of the Royal Society B: Biological Sciences*, **268**(1462): 45–52.

Dye, C (2009) Doomsday postponed? Preventing and reversing epidemics of drug-resistant tuberculosis. *Nature Reviews Microbiology*, **7**(1): 81–87.

Dye, C, Williams, BG, Espinal, MA, Raviglione, MC (2002) Erasing the world's slow stain: strategies to beat multidrug-resistant tuberculosis. *Science (New York, N.Y.)*, **295**(5562): 2042–6.

Egger, M, Ekouevi, DK, Williams, C, Lyamuya, RE, Mukumbi, H, Braitstein, P, Hartwell, T, Graber, C, Chi, BH, Boulle, A, Dabis, F, Wools-Kaloustian, K (2012) Cohort Profile: The international epidemiological databases to evaluate AIDS (IeDEA) in sub-Saharan Africa. *International Journal of Epidemiology*, **41**(5): 1256–1264.

Eldholm, V, Monteserin, J, Rieux, A, Lopez, B, Sobkowiak, B, Ritacco, V, Balloux, F (2015) Four decades of transmission of a multidrug-resistant Mycobacterium tuberculosis outbreak strain. en. *Nature Communications*, **6** 7119.

Eldholm, V, Norheim, G, Lippe, B von der, Kinander, W, Dahle, UR, Caugant, Da, Mannsåker, T, Mengshoel, AT, Dyrhol-Riise, AM, Balloux, F (2014) Evolution of extensively drug-resistant Mycobacterium tuberculosisfrom a susceptible ancestor in a single patient. *Genome Biology*, **15** 490.

Engström, A (2016) Fighting an old disease with modern tools: characteristics and molecular detection methods of drug-resistant Mycobacterium tuberculosis. *Infectious Diseases*, **48**(1): 1–17.

Espinal, MA, Dye, C (2005) Can DOTS control multidrug-resistant tuberculosis? *The Lancet*, **365**(9466): 1206–1209.

Farhat, MR, Jacobson, KR, Franke, MF, Kaur, D, Sloutsky, A, Mitnick, CD, Murray, M (2016) Gyrase Mutations Are Associated with Variable Levels of Fluoroquinolone Resistance in Mycobacterium tuberculosis. *Journal of Clinical Microbiology*, **54**(3): 727–733.

Farhat, MR, Shapiro, BJ, Kieser, KJ, Sultana, R, Jacobson, KR, Victor, TC, Warren, RM, Streicher, EM, Calver, A, Sloutsky, A, Kaur, D, Posey, JE, Plikaytis, B, Oggioni, MR, Gardy, JL, Johnston, JC, Rodrigues, M, Tang, PKC, Kato-Maeda, M, Borowsky, ML, Muddukrishna, B, Kreiswirth, BN, Kurepina, N, Galagan, J, Gagneux, S, Birren, B, Rubin, EJ, Lander, ES, Sabeti, PC, Murray, M (2013) Genomic analysis identifies targets of convergent positive selection in drug-resistant Mycobacterium tuberculosis. *Nature genetics*, **45**(10): 1183–9.

Feng-Zeng, Z, Murray, C, Spinaci, S, Styblo, K, Broekmans, J, Ji-Qian, Z, Yuang-Sheng, D, Yan-Hua, C, Xin, D, Ji-Ping, C, Yan-Lin, C, Li-Mao, L, Ning-Wu, L, Chuan-Jin, L, Hong-Qiao, L, Qiang, L, Da-Rang, T, Peng-Fei, W, Zhen-Han, W, Pu-Er, W, Fa-Da, Z, Fu-Sheng, Z, Gui-Lin, Z (1996) Results of directly observed short-course chemotherapy

in 112 842 Chinese patients with smear-positive tuberculosis. *The Lancet*, **347**(8998): 358–362.

Fenner, L, Egger, M, Bodmer, T, Altpeter, E, Zwahlen, M, Jaton, K, Pfyffer, GE, Borrell, S, Dubuis, O, Bruderer, T, Siegrist, HH, Furrer, H, Calmy, A, Fehr, J, Stalder, JM, Ninet, B, Böttger, EC, Gagneux, S (2012) Effect of Mutation and Genetic Background on Drug Resistance in Mycobacterium tuberculosis. *Antimicrobial Agents and Chemotherapy*, **56**(6): 3047–3053.

Ford, CB, Lin, PL, Chase, MR, Shah, RR, Iartchouk, O, Galagan, J, Mohaideen, N, Ioerger, TR, Sacchettini, JC, Lipsitch, M, Flynn, JL, Fortune, SM (2011) Use of whole genome sequencing to estimate the mutation rate of Mycobacterium tuberculosis during latent infection. *Nature Genetics*, **43**(5): 482–486.

Ford, CB, Shah, RR, Maeda, MK, Gagneux, S, Murray, MB, Cohen, T, Johnston, JC, Gardy, J, Lipsitch, M, Fortune, SM (2013) Mycobacterium tuberculosis mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis. *Nature Genetics*, **45**(7): 784–790.

Frank, SA (2010) The trade-off between rate and yield in the design of microbial metabolism. *Journal of evolutionary biology*, **23**(3): 609–13.

Franzblau, SG, DeGroote, MA, Cho, SH, Andries, K, Nuermberger, E, Orme, IM, Mdluli, K, Angulo-Barturen, I, Dick, T, Dartois, V, Lenaerts, AJ (2012) Comprehensive analysis of methods used for the evaluation of compounds against Mycobacterium tuberculosis. *Tuberculosis (Edinburgh, Scotland)*, **92**(6): 453–88.

Freihofer, P, Akbergenov, R, Teo, Y, Juskeviciene, R, Andersson, DI, Böttger, EC (2016) Nonmutational compensation of the fitness cost of antibiotic resistance in mycobacteria by overexpression of tlyA rRNA methylase. *RNA*, **22**(12): 1836–1843.

Frieden, TR, Fujiwara, PI, Washko, RM, Hamburg, MA (1995) Tuberculosis in New York City — Turning the Tide. *New England Journal of Medicine*, **333**(4): 229–233.

Frieden, TR, Sbarbaro, JA (2007) Promoting adherence to treatment for tuberculosis: the importance of direct observation. *Bulletin of the World Health Organization*, **85**(5): 407–9.

Gagneux, S, Burgos, MV, DeRiemer, K, Encisco, A, Muñoz, S, Hopewell, PC, Small, PM, Pym, AS (2006a) Impact of bacterial genetics on the transmission of isoniazid-resistant Mycobacterium tuberculosis. *PLoS pathogens*, **2**(6): e61.

Gagneux, S (2018) Ecology and evolution of Mycobacterium tuberculosis. *Nature Reviews Microbiology*, **16**(4): 202–213.

Gagneux, S, Long, CD, Small, PM, Van, T, Schoolnik, GK, Bohannan, BJM (2006b) The competitive cost of antibiotic resistance in Mycobacterium tuberculosis. *Science (New York, N.Y.)*, **312**(5782): 1944–6.

Gagneux, S, Long, CD, Small, PM, Van, T, Schoolnik, GK, Bohannan, BJM (2006c) The competitive cost of antibiotic resistance in Mycobacterium tuberculosis. Supplement. *Science (New York, N.Y.)*, **312**(5782): 1944–6.

Gagneux, S, Small, PM (2007) Global phylogeography of Mycobacterium tuberculosis and implications for tuberculosis product development. *The Lancet infectious diseases*, **7**(5): 328–37.

Gandhi, NR, Moll, A, Sturm, aW, Pawinski, R, Govender, T, Lalloo, U, Zeller, K, Andrews, J, Friedland, G (2006) Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa. *The Lancet*, **368**(9547): 1575–1580.

Ghielmetti, G, Coscolla, M, Ruetten, M, Friedel, U, Loiseau, C, Feldmann, J, Steinmetz, HW, Stucki, D, Gagneux, S (2017) Tuberculosis in Swiss captive Asian elephants: microevolution of Mycobacterium tuberculosis characterized by multilocus variable-number tandem-repeat analysis and whole-genome sequencing. *Scientific Reports*, **7**(1): 14647.

Gillespie, SH, Basu, S, Dickens, AL, O'Sullivan, DM, McHugh, TD (2005) Effect of subinhibitory concentrations of ciprofloxacin on Mycobacterium fortuitum mutation rates. *The Journal of antimicrobial chemotherapy*, **56**(2): 344–8.

Griffiths, AJ, Wessler, SR, Lewontin, RC, Gelbart, WM, Suzuki, DT, H., MJ (2005) *An introduction to genetic analysis*. 8th. New York: W H Freeman & Co, 706.

Gullberg, E, Cao, S, Berg, OG, Ilbäck, C, Sandegren, L, Hughes, D, Andersson, DI (2011) Selection of Resistant Bacteria at Very Low Antibiotic Concentrations. *PLoS Pathogens*, **7**(7): ed. by M Lipsitch, e1002158.

Gumbo, T (2010) New Susceptibility Breakpoints for First-Line Antituberculosis Drugs Based on Antimicrobial Pharmacokinetic/Pharmacodynamic Science and Population Pharmacokinetic Variability. *Antimicrobial Agents and Chemotherapy*, **54**(4): 1484–1491.

Gumbo, T, Louie, A, Liu, W, Brown, D, Ambrose, PG, Bhavnani, SM, Drusano, GL (2007) Isoniazid bactericidal activity and resistance emergence: Integrating pharmacodynamics and pharmacogenomics to predict efficacy in different ethnic populations. *Antimicrobial Agents and Chemotherapy*, **51**(7): 2329–2336.

Gupta, AK, Katoch, VM, Chauhan, DS, Sharma, R, Singh, M, Venkatesan, K, Sharma, VD (2010) Microarray Analysis of Efflux Pump Genes in Multidrug-Resistant My-

cobacterium tuberculosis During Stress Induced by Common Anti-Tuberculous Drugs. *Microbial Drug Resistance*, **16**(1): 21–28.

Gurbanova, E, Mehdiyev, R, Blondal, K, Altraja, A (2016) Predictors of cure in rifampicin-resistant tuberculosis in prison settings with low loss to follow-up. *The International Journal of Tuberculosis and Lung Disease*, **20**(5): 645–651.

Gygli, SM, Borrell, S, Trauner, A, Gagneux, S (2017) Antimicrobial resistance in Mycobacterium tuberculosis: mechanistic and evolutionary perspectives. *FEMS Microbiology Reviews*, **41**(3): 354–373.

Gygli, SM, Keller, PM, Ballif, M, Blöchliger, N, Hömke, R, Reinhard, M, Loiseau, C, Ritter, C, Sander, P, Borrell, S, Loo, JC, Avihingsanon, A, Gnokoro, J, Yotebieng, M, Egger, M, Gagneux, S, Böttger, EC (2018) Whole genome sequencing for drug resistance profile prediction in Mycobacterium tuberculosis. *bioRxiv*, (401703):

Hall, AR, Angst, DC, Schiessl, KT, Ackermann, M (2015) Costs of antibiotic resistance - separating trait effects and selective effects. *Evolutionary applications*, **8**(3): 261–72.

Hartkoorn, RC, Uplekar, S, Cole, ST (2014) Cross-resistance between clofazimine and bedaquiline through upregulation of MmpL5 in Mycobacterium tuberculosis. *Antimicrobial agents and chemotherapy*, **58**(5): 2979–81.

Hatherell, HA, Colijn, C, Stagg, HR, Jackson, C, Winter, JR, Abubakar, I (2016) Interpreting whole genome sequencing for investigating tuberculosis transmission: a systematic review. *BMC medicine*, **14**(1): 21.

Haver, HL, Chua, A, Ghode, P, Lakshminarayana, SB, Singhal, A, Mathema, B, Wintjens, R, Bifani, P (2015) Mutations in genes for the F420 biosynthetic pathway and a nitroreductase enzyme are the primary resistance determinants in spontaneous in vitro-selected PA-824-resistant mutants of Mycobacterium tuberculosis. *Antimicrobial agents and chemotherapy*, **59**(9): 5316–23.

Hazbón, MH, Brimacombe, M, Bobadilla del Valle, M, Cavatore, M, Guerrero, MI, Varma-Basil, M, Billman-Jacobe, H, Lavender, C, Fyfe, J, García-García, L, León, CI, Bose, M, Chaves, F, Murray, M, Eisenach, KD, Sifuentes-Osornio, J, Cave, MD, Ponce de León, A, Alland, D (2006) Population genetics study of isoniazid resistance mutations and evolution of multidrug-resistant Mycobacterium tuberculosis. *Antimicrobial agents and chemotherapy*, **50**(8): 2640–9.

Helaine, S, Holden, DW, Mouton, JM, Sampson, SL (2016) Elucidating population-wide mycobacterial replication dynamics at the single-cell level. *Microbiology*, **162**(6): 966–978.

Hershberg, R, Lipatov, M, Small, PM, Sheffer, H, Niemann, S, Homolka, S, Roach, JC, Kremer, K, Petrov, Da, Feldman, MW, Gagneux, S (2008) High Functional Diversity in

Mycobacterium tuberculosis Driven by Genetic Drift and Human Demography. *PLoS Biology*, **6**(12): ed. by MJ Blaser, e311.

Heym, B, Alzari, PM, Honore, N, Cole, ST (1995) Missense mutations in the catalase-peroxidase gene, katG, are associated with isoniazid resistance in Mycobacterium tuberculosis. *Molecular Microbiology*, **15**(2): 235–245.

Heym, B, Stavropoulos, E, Honoré, N, Domenech, P, Saint-Joanis, B, Wilson, TM, Collins, DM, Colston, MJ, Cole, ST (1997) Effects of overexpression of the alkyl hydroperoxide reductase AhpC on the virulence and isoniazid resistance of Mycobacterium tuberculosis. *Infection and immunity*, **65**(4): 1395–401.

Hillemann, D, Rusch-Gerdes, S, Richter, E (2008) In Vitro-Selected Linezolid-Resistant Mycobacterium tuberculosis Mutants. *Antimicrobial Agents and Chemotherapy*, **52**(2): 800–801.

Houghton, JL, Green, KD, Pricer, RE, Mayhoub, AS, Garneau-Tsodikova, S (2013) Unexpected N-acetylation of capreomycin by mycobacterial Eis enzymes. *Journal of Antimicrobial Chemotherapy*, **68**(4): 800–805.

Hughes, D, Brandis, G (2013) Rifampicin Resistance: Fitness Costs and the Significance of Compensatory Evolution. *Antibiotics*, **2**(2): 206–216.

Hugonnet, JE, Blanchard, JS (2007) Irreversible inhibition of the Mycobacterium tuberculosis beta-lactamase by clavulanate. *Biochemistry*, **46**(43): 11998–2004.

Hugonnet, JE, Tremblay, LW, Boshoff, HI, Barry, CE, Blanchard, JS (2009) Meropenem-Clavulanate Is Effective Against Extensively Drug-Resistant Mycobacterium tuberculosis. *Science*, **323**(5918): 1215–1218. eprint: `1002.1037`.

Huitric, E, Verhasselt, P, Koul, A, Andries, K, Hoffner, S, Andersson, DI (2010) Rates and Mechanisms of Resistance Development in Mycobacterium tuberculosis to a Novel Diarylquinoline ATP Synthase Inhibitor. *Antimicrobial Agents and Chemotherapy*, **54**(3): 1022–1028.

Ismail, NA, Omar, SV, Joseph, L, Govender, N, Blows, L, Ismail, F, Koornhof, H, Dreyer, AW, Kaniga, K, Ndjeka, N (2018) Defining Bedaquiline Susceptibility, Resistance, Cross-Resistance and Associated Genetic Determinants: A Retrospective Cohort Study. *EBioMedicine*, **28** 136–142.

Jackson, M (2014) The Mycobacterial Cell Envelope–Lipids. *Cold Spring Harbor Perspectives in Medicine*, **4**(10): a021105–a021105.

Jacobs, A (1941) Infective dose in pulmonary tuberculosis. *Tubercle*, **22**(11): 266–271.

Jagannath, C, Reddy, VM, Gangadharam, PRJ (1995) Enhancement of drug susceptibility of multi-drug resistant strains of Mycobacterium tuberculosis by ethambutol and dimethyl sulphoxide. *Journal of Antimicrobial Chemotherapy*, **35**(3): 381–390.

Janardhan, S, Ram Vivek, M, Narahari Sastry, G (2016) Modeling the permeability of drug-like molecules through the cell wall of Mycobacterium tuberculosis: an analogue based approach. *Mol. BioSyst.*, **12**(11): 3377–3384.

Jarlier, V, Nikaido, H (1994) Mycobacterial cell wall: structure and role in natural resistance to antibiotics. *FEMS microbiology letters*, **123** 11–18.

Jassal, MS, Bishai, WR (2010) Epidemiology and Challenges to the Elimination of Global Tuberculosis. *Clinical Infectious Diseases*, **50**(s3): S156–S164. eprint: 15334406.

Jin, DJ, Gross, CA (1991) RpoB8, a rifampicin-resistant termination-proficient RNA polymerase, has an increased Km for purine nucleotides during transcription elongation. *The Journal of biological chemistry*, **266**(22): 14478–85.

Kambli, P, Ajbani, K, Nikam, C, Sadani, M, Shetty, A, Udwadia, Z, Georghiou, SB, Rodwell, TC, Catanzaro, A, Rodrigues, C (2016) Correlating rrs and eis promoter mutations in clinical isolates of Mycobacterium tuberculosis with phenotypic susceptibility levels to the second-line injectables. *International Journal of Mycobacteriology*, **5**(1): 1–6.

Kaplan, G, Post, FA, Moreira, AL, Wainwright, H, Kreiswirth, BN, Tanverdi, M, Mathema, B, Ramaswamy, SV, Walther, G, Steyn, LM, Barry, CE, Bekker, LG (2003a) Mycobacterium tuberculosis Growth at the Cavity Surface: a Microenvironment with Failed Immunity. *Infection and Immunity*, **71**(12): 7099–7108.

Kaplan, G, Post, FA, Moreira, AL, Wainwright, H, Kreiswirth, BN, Tanverdi, M, Mathema, B, Ramaswamy, SV, Walther, G, Steyn, LM, Barry, CE, Bekker, LG (2003b) Mycobacterium tuberculosis Growth at the Cavity Surface: a Microenvironment with Failed Immunity. *Infection and Immunity*, **71**(12): 7099–7108.

Kaufmann, SHE, Lange, C, Rao, M, Balaji, KN, Lotze, M, Schito, M, Zumla, AI, Maeurer, M (2014) Progress in tuberculosis vaccine development and host-directed therapies—a state of the art review. *The Lancet Respiratory Medicine*, **2**(4): 301–320.

Kenyon, G (2009) Russia's prisons fuel drug-resistant tuberculosis. *The Lancet Infectious Diseases*, **9**(10): 594.

Knight, GM, Colijn, C, Shrestha, S, Fofana, M, Cobelens, F, White, RG, Dowdy, DW, Cohen, T (2015) The Distribution of Fitness Costs of Resistance-Conferring Mutations Is a Key Determinant for the Future Burden of Drug-Resistant Tuberculosis: A Model-Based Analysis. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, **61Suppl 3**(suppl 3): S147–54.

Koch, A, Mizrahi, V, Warner, DF (2014) The impact of drug resistance on Mycobacterium tuberculosis physiology: what can we learn from rifampicin? *Emerging Microbes & Infections*, **3**(3): e17.

Koch, R (1882) Die Ätiologie der Tuberkulose. *Berliner klinische Wochenschrift*, **19**(15): 221–30.

Konno, K, Feldmann, FM, McDermott, W (1967) Pyrazinamide susceptibility and amidase activity of tubercle bacilli. *The American review of respiratory disease*, **95**(3): 461–9.

Kranzer, K, Afnan-Holmes, H, Tomlin, K, Golub, JE, Shapiro, AE, Schaap, A, Corbett, EL, Lönnroth, K, Glynn, JR (2013) The benefits to communities and individuals of screening for active tuberculosis disease: A systematic review. *International Journal of Tuberculosis and Lung Disease*, **17**(4): 432–446. eprint: `/dx.doi.org/10.5588/ijtld.12.0743`.

Lamichhane, G, Tyagi, S, Bishai, WR (2005) Designer Arrays for Defined Mutant Analysis To Detect Genes Essential for Survival of Mycobacterium tuberculosis in Mouse Lungs. *Infection and Immunity*, **73**(4): 2533–2540.

Lanzas, F, Karakousis, PC, Sacchettini, JC, Ioerger, TR (2013) Multidrug-resistant tuberculosis in panama is driven by clonal expansion of a multidrug-resistant Mycobacterium tuberculosis strain related to the KZN extensively drug-resistant M. tuberculosis strain from South Africa. *Journal of clinical microbiology*, **51**(10): 3277–85.

Law, S, Piatek, AS, Vincent, C, Oxlade, O, Menzies, D (2017) Emergence of drug resistance in patients with tuberculosis cared for by the Indian health-care system: a dynamic modelling study. *The Lancet Public Health*, **2**(1): e47–e55.

Lee, J, Armstrong, DT, Ssengooba, W, Park, JA, Yu, Y, Mumbowa, F, Namaganda, C, Mboowa, G, Nakayita, G, Armakovitch, S, Chien, G, Cho, SN, Via, LE, Barry, CE, Ellner, JJ, Alland, D, Dorman, SE, Joloba, ML (2014) Sensititre MYCOTB MIC Plate for Testing Mycobacterium tuberculosis Susceptibility to First- and Second-Line Drugs. *Antimicrobial Agents and Chemotherapy*, **58**(1): 11–18.

Lehmann, J (1946) PARA-AMINOSALICYLIC ACID IN THE TREATMENT OF TUBERCULOSIS. *The Lancet*, **247**(6384): 15–16.

Li, G, Zhang, J, Guo, Q, Jiang, Y, Wei, J, Zhao, Ll, Zhao, X, Lu, J, Wan, K (2015) Efflux Pump Gene Expression in Multidrug-Resistant Mycobacterium tuberculosis Clinical Isolates. *PLOS ONE*, **10**(2): ed. by D Chatterji, e0119013.

Lieberman, TD, Wilson, D, Misra, R, Xiong, LL, Moodley, P, Cohen, T, Kishony, R (2016) Genomic diversity in autopsy samples reveals within-host dissemination of HIV-associated Mycobacterium tuberculosis. *Nature Medicine*, **22**(12): 1470–1474.

Lin, N, Liu, Z, Zhou, J, Wang, S, Fleming, J (2013) Draft Genome Sequences of Two Super-XDR Isolates of M. tuberculosis from China. *FEMS microbiology letters*, **2** 2–5.

Lin, PL, Rodgers, M, Smith, L, Bigbee, M, Myers, A, Bigbee, C, Chiosea, I, Capuano, SV, Fuhrman, C, Klein, E, Flynn, JL (2009) Quantitative Comparison of Active and Latent Tuberculosis in the Cynomolgus Macaque Model. *Infection and Immunity*, **77**(10): 4631–4642.

Liu, J, Nikaido, H (1999) A mutant of Mycobacterium smegmatis defective in the biosynthesis of mycolic acids accumulates meromycolates. *Proceedings of the National Academy of Sciences*, **96**(7): 4011–4016.

Liu, Q, Zuo, T, Xu, P, Jiang, Q, Wu, J, Gan, M, Yang, C, Prakash, R, Zhu, G, Takiff, HE, Gao, Q (2018) Have compensatory mutations facilitated the current epidemic of multidrug-resistant tuberculosis? *Emerging Microbes & Infections*, **7**(1): 1–8.

Lomtadze, N, Aspindzelashvili, R, Janjgava, M, Mirtskhulava, V, Wright, A, Blumberg, HM, Salakaia, A (2009) Prevalence and risk factors for multidrug-resistant tuberculosis in the Republic of Georgia: a population-based study. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease*, **13**(1): 68–73. eprint: NIHMS150003.

Louw, GE, Warren, RM, Gey van Pittius, NC, McEvoy, CRE, Van Helden, PD, Victor, TC (2009) A Balancing Act: Efflux/Influx in Mycobacterial Drug Resistance. *Antimicrobial Agents and Chemotherapy*, **53**(8): 3181–3189.

Luciani, F, Sisson, Sa, Jiang, H, Francis, AR, Tanaka, MM (2009) The epidemiological fitness cost of drug resistance in Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences*, **106**(34): 14711–14715.

Luo, T, Yang, C, Peng, Y, Lu, L, Sun, G, Wu, J, Jin, X, Hong, J, Li, F, Mei, J, Deriemer, K, Gao, Q (2014) Whole-genome sequencing to detect recent transmission of Mycobacterium tuberculosis in settings with a high burden of tuberculosis. *Tuberculosis*, **94**(4): 434–440.

Luria, SE, Delbrück, M (1943) Mutations of Bacteria from Virus Sensitivity to Virus Resistance. *Genetics*, **28**(6): 491–511.

Mabud, TS, de Lourdes Delgado Alves, M, Ko, AI, Basu, S, Walter, KS, Cohen, T, Mathema, B, Colijn, C, Lemos, E, Croda, J, Andrews, JR (2019) Evaluating strategies for control of tuberculosis in prisons and prevention of spillover into communities: An observational and modeling study from Brazil. *PLOS Medicine*, **16**(1): ed. by JZ Metcalfe, e1002737.

Machado, D, Coelho, TS, Perdigão, J, Pereira, C, Couto, I, Portugal, I, Maschmann, RDA, Ramos, DF, Groll, A von, Rossetti, ML, Silva, PA, Viveiros, M (2017) Interplay between mutations and efflux in drug resistant clinical isolates of Mycobacterium tuberculosis. *Frontiers in Microbiology*, **8**(APR): 1–18.

Machado, D, Couto, I, Perdigão, J, Rodrigues, L, Portugal, I, Baptista, P, Veigas, B, Amaral, L, Viveiros, M (2012) Contribution of Efflux to the Emergence of Isoniazid and Multidrug Resistance in Mycobacterium tuberculosis. *PLoS ONE*, **7**(4): ed. by I Mokrousov, e34538.

Madsen, CT, Jakobsen, L, Buriánková, K, Doucet-Populaire, F, Pernodet, JL, Douthwaite, S (2005) Methyltransferase Erm(37) slips on rRNA to confer atypical resistance in Mycobacterium tuberculosis. *The Journal of biological chemistry*, **280**(47): 38942–7.

Mahmoudi, A, Iseman, MD (1993) Pitfalls in the care of patients with tuberculosis. Common errors and their association with the acquisition of drug resistance. *JAMA*, **270**(1): 65–8.

Mailaender, C, Reiling, N, Engelhardt, H, Bossmann, S, Ehlers, S, Niederweis, M (2004) The MspA porin promotes growth and increases antibiotic susceptibility of both Mycobacterium bovis BCG and Mycobacterium tuberculosis. *Microbiology*, **150**(4): 853–864.

Maisnier-Patin, S, Berg, OG, Liljas, L, Andersson, DI (2002) Compensatory adaptation to the deleterious effect of antibiotic resistance in Salmonella typhimurium. *Molecular Microbiology*, **46**(2): 355–366.

Malik, S, Willby, M, Sikes, D, Tsodikov, OV, Posey, JE (2012) New Insights into Fluoroquinolone Resistance in Mycobacterium tuberculosis: Functional Genetic Analysis of gyrA and gyrB Mutations. *PLoS ONE*, **7**(6): e39754.

Manca, C, Paul, S, Barry, CE, Freedman, VH, Kaplan, G (1999) Mycobacterium tuberculosis catalase and peroxidase activities and resistance to oxidative killing in human monocytes in vitro. *Infection and immunity*, **67**(1): 74–9.

Manjunatha, UH, Boshoff, H, Dowd, CS, Zhang, L, Albert, TJ, Norton, JE, Daniels, L, Dick, T, Pang, SS, Barry, CE (2006) Identification of a nitroimidazo-oxazine-specific protein involved in PA-824 resistance in Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences of the United States of America*, **103**(2): 431–436.

Martin Bland, J, Altman, D (1986) STATISTICAL METHODS FOR ASSESSING AGREEMENT BETWEEN TWO METHODS OF CLINICAL MEASUREMENT. *The Lancet*, **327**(8476): 307–310.

Maruri, F, Sterling, TR, Kaiga, AW, Blackman, A, Heijden, YF van der, Mayer, C, Cambau, E, Aubry, A (2012) A systematic review of gyrase mutations associated with fluoroquinolone-resistant Mycobacterium tuberculosis and a proposed gyrase numbering system. *Journal of Antimicrobial Chemotherapy*, **67**(4): 819–831.

Maus, CE, Plikaytis, BB, Shinnick, TM (2005) Molecular Analysis of Cross-Resistance to Capreomycin, Kanamycin, Amikacin, and Viomycin in Mycobacterium tuberculosis. *Antimicrobial Agents and Chemotherapy*, **49**(8): 3192–3197.

Mazando, S, Zimudzi, C, Zimba, M, Sande, S, Gundidza, M, Mazorodze, JH, Seepe, PMM, Pym, A, Mason, P (2017) High efflux pump activity and gene expression at baseline linked to poor tuberculosis treatment outcomes. *Journal of Medical and Biomedical Sciences*, **6** 8–17.

McGrath, M, Gey van Pittius, NC, Helden, PD van, Warren, RM, Warner, DF (2014) Mutation rate and the emergence of drug resistance in Mycobacterium tuberculosis. *Journal of Antimicrobial Chemotherapy*, **69**(2): 292–302.

Medical Research Council (1950) Treatment of Pulmonary Tuberculosis with Streptomycin and Para-Amino-Salicylic Acid: A Medical Research Council Investigation. *BMJ*, **2**(4688): 1073–1085.

Meftahi, N, Namouchi, A, Mhenni, B, Brandis, G, Hughes, D, Mardassi, H (2016) Evidence for the critical role of a secondary site rpoB mutation in the compensatory evolution and successful transmission of an MDR tuberculosis outbreak strain. *Journal of Antimicrobial Chemotherapy*, **71**(2): 324–332.

Meier, A, Kirschner, P, Bange, FC, Vogel, U, Böttger, EC (1994) Genetic alterations in streptomycin-resistant Mycobacterium tuberculosis: mapping of mutations conferring resistance. *Antimicrobial agents and chemotherapy*, **38**(2): 228–33.

Melnyk, AH, Wong, A, Kassen, R (2015) The fitness costs of antibiotic resistance mutations. *Evolutionary Applications*, **8**(3): 273–283.

Merker, M, Barbier, M, Cox, H, Rasigade, JP, Feuerriegel, S, Kohl, TA, Diel, R, Borrell, S, Gagneux, S, Nikolayevskyy, V, Andres, S, Nübel, U, Supply, P, Wirth, T, Niemann, S (2018) Compensatory evolution drives multidrug-resistant tuberculosis in Central Asia. *eLife*, **7** 1–31.

Merker, M, Blin, C, Mona, S, Duforet-Frebourg, N, Lecher, S, Willery, E, Blum, MGB, Rüsch-Gerdes, S, Mokrousov, I, Aleksic, E, Allix-Béguec, C, Antierens, A, Augustynowicz-Kopeć, E, Ballif, M, Barletta, F, Beck, HP, Barry, CE, Bonnet, M, Borroni, E, Campos-Herrero, I, Cirillo, D, Cox, H, Crowe, S, Crudu, V, Diel, R, Drobniewski, F, Fauville-Dufaux, M, Gagneux, S, Ghebremichael, S, Hanekom, M, Hoffner, S, Jiao, Ww, Kalon, S, Kohl, TA, Kontsevaya, I, Lillebæk, T, Maeda, S, Nikolayevskyy, V, Rasmussen, M, Rastogi, N, Samper, S, Sanchez-Padilla, E, Savic, B, Shamputa, IC, Shen, A, Sng, LH, Stakenas, P, Toit, K, Varaine, F, Vukovic, D, Wahl, C, Warren, R, Supply, P, Niemann, S, Wirth, T (2015) Evolutionary history and global spread of the Mycobacterium tuberculosis Beijing lineage. *Nature Genetics*, **47**(3): 242–249.

Merker, M, Kohl, Ta, Roetzer, A, Truebe, L, Richter, E, Rüsch-Gerdes, S, Fattorini, L, Oggioni, MR, Cox, H, Varaine, F, Niemann, S (2013) Whole Genome Sequencing Reveals Complex Evolution Patterns of Multidrug-Resistant Mycobacterium tuberculosis Beijing Strains in Patients. *PLoS ONE*, **8**(12): ed. by JZ Metcalfe, e82551.

Mestre, O, Luo, T, Dos Vultos, T, Kremer, K, Murray, A, Namouchi, A, Jackson, C, Rauzier, J, Bifani, P, Warren, R, Rasolofo, V, Mei, J, Gao, Q, Gicquel, B (2011) Phylogeny of Mycobacterium tuberculosis Beijing Strains Constructed from Polymorphisms in Genes Involved in DNA Replication, Recombination and Repair. *PLoS ONE*, **6**(1): ed. by AK Tyagi, e16020.

Middlebrook, G, Cohn, ML (1953) Some observations on the pathogenicity of isoniazid-resistant variants of tubercle bacilli. *Science (New York, N.Y.)*, **118**(3063): 297–9.

Milano, A, Pasca, MR, Provvedi, R, Lucarelli, AP, Manina, G, Luisa de Jesus Lopes Ribeiro, A, Manganelli, R, Riccardi, G (2009) Azole resistance in Mycobacterium tuberculosis is mediated by the MmpS5–MmpL5 efflux system. *Tuberculosis*, **89**(1): 84–90.

Minato, Y, Thiede, JM, Kordus, SL, McKlveen, EJ, Turman, BJ, Baughn, AD (2015) Mycobacterium tuberculosis folate metabolism and the mechanistic basis for para-aminosalicylic acid susceptibility and resistance. *Antimicrobial Agents and Chemotherapy*, **59**(9): 5097–5106.

Miotto, P, Cabibbe, AM, Borroni, E, Degano, M, Cirillo, DM (2018) Role of Disputed Mutations in the rpoB Gene in Interpretation of Automated Liquid MGIT Culture Results for Rifampin Susceptibility Testing of Mycobacterium tuberculosis. *Journal of Clinical Microbiology*, **56**(5): ed. by GA Land, JCM.01599–17.

Miotto, P, Tessema, B, Tagliani, E, Chindelevitch, L, Starks, AM, Emerson, C, Hanna, D, Kim, PS, Liwski, R, Zignol, M, Gilpin, C, Niemann, S, Denkinger, CM, Fleming, J, Warren, RM, Crook, D, Posey, J, Gagneux, S, Hoffner, S, Rodrigues, C, Comas, I, Engelthaler, DM, Murray, M, Alland, D, Rigouts, L, Lange, C, Dheda, K, Hasan, R, Ranganathan, UDK, McNerney, R, Ezewudo, M, Cirillo, DM, Schito, M, Köser, CU, Rodwell, TC (2017) A standardised method for interpreting the association between mutations and phenotypic drug resistance in <>Mycobacterium tuberculosis<>. *European Respiratory Journal*, **50**(6): 1701354.

Mokrousov, I, Vyazovaya, A, Otten, T, Zhuravlev, V, Pavlova, E, Tarashkevich, L, Krishevich, V, Vishnevsky, B, Narvskaya, O (2012) Mycobacterium tuberculosis Population in Northwestern Russia: An Update from Russian-EU/Latvian Border Region. *PLoS ONE*, **7**(7): ed. by PCY Woo, e41318.

Monshupanee, T, Johansen, SK, Dahlberg, AE, Douthwaite, S (2012) Capreomycin susceptibility is increased by TlyA-directed 2- O -methylation on both ribosomal subunits. *Molecular Microbiology*, **85**(6): 1194–1203.

Moreno-Gamez, S, Hill, AL, Rosenbloom, DIS, Petrov, DA, Nowak, MA, Pennings, PS (2015) Imperfect drug penetration leads to spatial monotherapy and rapid evolution of multidrug resistance. *Proceedings of the National Academy of Sciences*, **112**(22): E2874–E2883.

Morlock, GP, Metchock, B, Sikes, D, Crawford, JT, Cooksey, RC (2003) ethA, inhA, and katG Loci of Ethionamide-Resistant Clinical Mycobacterium tuberculosis Isolates. *Antimicrobial Agents and Chemotherapy*, **47**(12): 3799–3805.

Moura de Sousa, J, Balbontín, R, Durão, P, Gordo, I (2017) Multidrug-resistant bacteria compensate for the epistasis between resistances. *PLOS Biology*, **15**(4): ed. by A de Visser, e2001741.

Movahedzadeh, F, Wheeler, PR, Dinadayala, P, Av-Gay, Y, Parish, T, Daffé, M, Stoker, NG (2010) Inositol monophosphate phosphatase genes of Mycobacterium tuberculosis. *BMC Microbiology*, **10**(1): 50.

Müller, B, Chihota, VN, Pillay, M, Klopper, M, Streicher, EM, Coetzee, G, Trollip, A, Hayes, C, Bosman, ME, Gey van Pittius, NC, Victor, TC, Gagneux, S, Helden, PD van, Warren, RM (2013) Programmatically Selected Multidrug-Resistant Strains Drive the Emergence of Extensively Drug-Resistant Tuberculosis in South Africa. *PLoS ONE*, **8**(8): ed. by PC Karakousis, e70919.

Munro, SA, Lewin, SA, Smith, HJ, Engel, ME, Fretheim, A, Volmink, J (2007) Patient Adherence to Tuberculosis Treatment: A Systematic Review of Qualitative Research. *PLoS Medicine*, **4**(7): ed. by B Rylko-Bauer, e238. eprint: 0208024 (gr-qc).

Musser, JM (1995) Antimicrobial agent resistance in mycobacteria: molecular genetic insights. *Clinical microbiology reviews*, **8**(4): 496–514.

Naidoo, C, Pillay, M (2014) Increased in vitro fitness of multi- and extensively drug-resistant F15/LAM4/KZN strains of Mycobacterium tuberculosis. *Clinical Microbiology and Infection*, **20**(6): O361–O369.

Nair, J, Rouse, DA, Bai, GH, Morris, SL (1993) The rpsL gene and streptomycin resistance in single and multiple drug-resistant strains of Mycobacterium tuberculosis. *Molecular Microbiology*, **10**(3): 521–527.

Nash, KA (2016) Multidrug Resistance in Mycobacteria. *Current Clinical Microbiology Reports*, **3**(1): 53–61.

Nathavitharana, RR, Hillemann, D, Schumacher, SG, Schlueter, B, Ismail, N, Omar, SV, Sikhondze, W, Havumaki, J, Valli, E, Boehme, C, Denkinger, CM (2016) Multicen-

ter Noninferiority Evaluation of Hain GenoType MTBDR plus Version 2 and Nipro NTM+MDRTB Line Probe Assays for Detection of Rifampin and Isoniazid Resistance. *Journal of Clinical Microbiology*, **54**(6): ed. by GA Land, 1624–1630.

Nguyen, L (2016) Antibiotic resistance mechanisms in M. tuberculosis: an update. *Archives of toxicology*, **90**(7): 1585–1604.

Nguyen, L, Pieters, J (2009) Mycobacterial subversion of chemotherapeutic reagents and host defense tactics: challenges in tuberculosis drug development. *Annual review of pharmacology and toxicology*, **49** 427–53.

Niemann, S, Diel, R, Khechinashvili, G, Gegia, M, Mdivani, N, Tang, YW (2010) Mycobacterium tuberculosis Beijing lineage favors the spread of multidrug-resistant tuberculosis in the Republic of Georgia. *Journal of clinical microbiology*, **48**(10): 3544–50.

O'Neill, J (2014) *Antimicrobial Resistance: Tackling a crisis for the health and wealth of nations*. Tech. rep.

Orhan, G, Bayram, A, Zer, Y, Balci, I (2005) Synergy tests by E test and checkerboard methods of antimicrobial combinations against Brucella melitensis. *Journal of clinical microbiology*, **43**(1): 140–3.

Pardini, M, Niemann, S, Varaine, F, Iona, E, Meacci, F, Orrù, G, Yesilkaya, H, Jarosz, T, Andrew, P, Barer, M, Checchi, F, Rinder, H, Orefici, G, Rüsch-Gerdes, S, Fattorini, L, Oggioni, MR, Bonnet, M (2009) Characteristics of drug-resistant tuberculosis in Abkhazia (Georgia), a high-prevalence area in Eastern Europe. *Tuberculosis*, **89**(4): 317–324.

Park, JS, Lee, JY, Lee, YJ, Kim, SJ, Cho, Y, Yoon, HI, Lee, CT, Song, J, Lee, JH (2016) Serum Levels of Antituberculosis Drugs and Their Effect on Tuberculosis Treatment Outcome. *Antimicrobial Agents and Chemotherapy*, **60**(1): 92–98.

Pasipanodya, JG, McIlleron, H, Burger, A, Wash, PA, Smith, P, Gumbo, T (2013) Serum drug concentrations predictive of pulmonary tuberculosis outcomes. *Journal of Infectious Diseases*, **208**(9): 1464–1473.

Paulsen, IT, Chen, J, Nelson, KE, Saier, MH (2001) Comparative genomics of microbial drug efflux systems. *Journal of molecular microbiology and biotechnology*, **3**(2): 145–50.

Paulson, T (2013) Epidemiology: A mortal foe. *Nature*, **502**(7470): S2–S3.

Payen, MC, De Wit, S, Martin, C, Sergysels, R, Muylle, I, Van Laethem, Y, Clumeck, N (2012) Clinical use of the meropenem-clavulanate combination for extensively drug-resistant tuberculosis. *The International Journal of Tuberculosis and Lung Disease*, **16**(4): 558–560.

Perkins, AE, Nicholson, WL (2008) Uncovering New Metabolic Capabilities of Bacillus subtilis Using Phenotype Profiling of Rifampin-Resistant rpoB Mutants. *Journal of Bacteriology*, **190**(3): 807–814.

Phelan, J, O'Sullivan, DM, Machado, D, Ramos, J, Whale, AS, O'Grady, J, Dheda, K, Campino, S, McNerney, R, Viveiros, M, Huggett, JF, Clark, TG (2016) The variability and reproducibility of whole genome sequencing technology for detecting resistance to anti-tuberculous drugs. *Genome Medicine*, **8**(1): 132.

Post, FA, Willcox, PA, Mathema, B, Steyn, LM, Shean, K, Ramaswamy, SV, Graviss, EA, Shashkina, E, Kreiswirth, BN, Kaplan, G (2004) Genetic Polymorphism in Mycobacterium tuberculosis Isolates from Patients with Chronic Multidrug-Resistant Tuberculosis. *The Journal of Infectious Diseases*, **190**(1): 99–106.

Prideaux, B, Via, LE, Zimmerman, MD, Eum, S, Sarathy, J, O'Brien, P, Chen, C, Kaya, F, Weiner, DM, Chen, PY, Song, T, Lee, M, Shim, TS, Cho, JS, Kim, W, Cho, SN, Olivier, KN, Barry, CE, Dartois, V (2015) The association between sterilizing activity and drug distribution into tuberculosis lesions. *Nature Medicine*, **21**(10): 1223–1227.

Pule, CM, Sampson, SL, Warren, RM, Black, PA, Helden, PD van, Victor, TC, Louw, GE (2016) Efflux pump inhibitors: targeting mycobacterial efflux systems to enhance TB therapy. *Journal of Antimicrobial Chemotherapy*, **71**(1): 17–26.

Pym, AS, Saint-Joanis, B, Cole, ST (2002) Effect of katG Mutations on the Virulence of Mycobacterium tuberculosis and the Implication for Transmission in Humans. *Infection and Immunity*, **70**(9): 4955–4960.

Qi, Q, Preston, GM, MacLean, RC (2014) Linking System-Wide Impacts of RNA Polymerase Mutations to the Fitness Cost of Rifampin Resistance in Pseudomonas aeruginosa. *mBio*, **5**(6): e01562–14.

Rabello, MCdS, Matsumoto, CK, Almeida, LGP de, Menendez, MC, Oliveira, RS de, Silva, RM, Garcia, MJ, Leão, SC (2012) First description of natural and experimental conjugation between Mycobacteria mediated by a linear plasmid. *PloS one*, **7**(1): 1–8.

Rancoita, PMV, Cugnata, F, Gibertoni Cruz, AL, Borroni, E, Hoosdally, SJ, Walker, TM, Grazian, C, Davies, TJ, Peto, TEA, Crook, DW, Fowler, PW, Cirillo, DM, CRyPTIC Consortiumfor the CRyPTIC Consortiumfor the CRyPTIC Consortium (2018) Validating a 14-Drug Microtiter Plate Containing Bedaquiline and Delamanid for Large-Scale Research Susceptibility Testing of Mycobacterium tuberculosis. *Antimicrobial agents and chemotherapy*, **62**(9): 1–15.

Rastogi, N, Goh, KS (1990) Action of 1-isonicotinyl-2-palmitoyl hydrazine against the Mycobacterium avium complex and enhancement of its activity by m-fluorophenylalanine. *Antimicrobial Agents and Chemotherapy*, **34**(11): 2061–2064.

Reichler, MR, Khan, A, Sterling, TR, Zhao, H, Moran, J, McAuley, J, Bessler, P, Mangura, B, Bakhtawar, I, LeDoux, C, McAuley, J, Beison, J, Fitzgerald, M, Naus, M, Nakajima, M, Schluger, N, Hirsch-Moverman, Y, Moran, J, Blumberg, H, Tapia, J, Singha, L, Hershfeld, E, Roche, B, Mangura, B, Sevilla, A, Sterling, T, Chavez-Lindell, T, Maruri, F, Dorman, S, Cronin, W, Munk, E, Khan, A, Yuan, Y, Chen, B, Yan, F, Shen, Y, Zhao, H, Zhang, H, Bessler, P, Fagley, M, Reichler, M, Reichler, M, Sterling, T, Tapia, J, Hirsch, C, Luo, C (2018) Risk and Timing of Tuberculosis Among Close Contacts of Persons with Infectious Tuberculosis. *The Journal of Infectious Diseases*, **218**

Reiling, N, Homolka, S, Walter, K, Brandenburg, J, Niwinski, L, Ernst, M, Herzmann, C, Lange, C, Diel, R, Ehlers, S, Niemann, S (2013) Clade-specific virulence patterns of Mycobacterium tuberculosis complex strains in human primary macrophages and aerogenically infected mice. *mBio*, **4**(4): e00250–13–e00250–13.

Reynolds, MG (2000) Compensatory evolution in rifampin-resistant Escherichia coli. *Genetics*, **156**(4): 1471–81.

Ritter, C, Lucke, K, Sirgel, FA, Warren, RW, Helden, PD van, Bottger, EC, Bloemberg, GV (2014) Evaluation of the AID TB Resistance Line Probe Assay for Rapid Detection of Genetic Alterations Associated with Drug Resistance in Mycobacterium tuberculosis Strains. *Journal of Clinical Microbiology*, **52**(3): 940–946.

Roa, MB, Tablizo, FA, Morado, EKD, Cunanan, LF, Uy, IDC, Ng, KCS, Manalastas-Cantos, KG, Reyes, JM, Ganchua, SKC, Ang, CF, Kato-Maeda, M, Cattamanchi, A, Karaoz, U, Destura, RV, Lluisma, AO (2018) Whole-genome sequencing and single nucleotide polymorphisms in multidrug-resistant clinical isolates of Mycobacterium tuberculosis from the Philippines. *Journal of Global Antimicrobial Resistance*, **15** 239–245.

Rock, JM, Lang, UF, Chase, MR, Ford, CB, Gerrick, ER, Gawande, R, Coscolla, M, Gagneux, S, Fortune, SM, Lamers, MH (2015) DNA replication fidelity in Mycobacterium tuberculosis is mediated by an ancestral prokaryotic proofreader. *Nature Genetics*, **47**(6): 677–681.

Rodriguez-Rivera, FP, Zhou, X, Theriot, JA, Bertozzi, CR (2017) Visualization of mycobacterial membrane dynamics in live cells. *Journal of the American Chemical Society*, **139**(9): 3488–3495.

Ruesen, C, Riza, AL, Florescu, A, Chaidir, L, Editoiu, C, Aalders, N, Nicolosu, D, Grecu, V, Ioana, M, Crevel, R van, Ingen, J van (2018) Linking minimum inhibitory concentrations to whole genome sequence-predicted drug resistance in Mycobacterium tuberculosis strains from Romania. *Scientific reports*, **8**(1): 9676.

Safi, H, Lingaraju, S, Amin, A, Kim, S, Jones, M, Holmes, M, McNeil, M, Peterson, SN, Chatterjee, D, Fleischmann, R, Alland, D (2013) Evolution of high-level ethambutol-resistant tuberculosis through interacting mutations in decaprenylphosphoryl-β-D-arabinose biosynthetic and utilization pathway genes. *Nature Genetics*, **45**(10): 1190–1197.

Sakamoto, K (2012) The Pathology of Mycobacterium tuberculosis Infection. *Veterinary Pathology*, **49**(3): 423–439.

Salvatore, PP, Becerra, MC, Abel zur Wiesch, P, Hinkley, T, Kaur, D, Sloutsky, A, Cohen, T (2016) Fitness Costs of Drug Resistance Mutations in Multidrug-Resistant Mycobacterium tuberculosis : A Household-Based Case-Control Study. *Journal of Infectious Diseases*, **213**(1): 149–155.

Sander, P, Springer, B, Prammananan, T, Sturmfels, A, Kappler, M, Pletschette, M, Böttger, EC (2002) Fitness cost of chromosomal drug resistance-conferring mutations. *Antimicrobial agents and chemotherapy*, **46**(5): 1204–11.

Sandgren, A, Strong, M, Muthukrishnan, P, Weiner, BK, Church, GM, Murray, MB (2009) Tuberculosis Drug Resistance Mutation Database. *PLoS Medicine*, **6**(2): e1000002.

Sarathy, JP, Dartois, V, Lee, EJD (2012) The role of transport mechanisms in Mycobacterium Tuberculosis drug resistance and tolerance. *Pharmaceuticals*, **5**(11): 1210–1235.

Schatz, A, Bugle, E, Waksman, SA (1944) Streptomycin, a Substance Exhibiting Antibiotic Activity Against Gram-Positive and Gram-Negative Bacteria. *Experimental Biology and Medicine*, **55**(1): 66–69.

Schmalstieg, AM, Srivastava, S, Belkaya, S, Deshpande, D, Meek, C, Leff, R, Oers, NSC van, Gumbo, T (2012) The antibiotic resistance arrow of time: efflux pump induction is a general first step in the evolution of mycobacterial drug resistance. *Antimicrobial agents and chemotherapy*, **56**(9): 4806–15.

Schön, T, Miotto, P, Köser, CU, Viveiros, M, Böttger, E, Cambau, E (2017) Mycobacterium tuberculosis drug-resistance testing: challenges, recent developments and perspectives. *Clinical Microbiology and Infection*, **23**(3): 154–160.

Scorpio, A, Zhang, Y (1996) Mutations in pncA, a gene encoding pyrazinamidase/nicotinamidase, cause resistance to the antituberculous drug pyrazinamide in tubercle bacillus. *Nature Medicine*, **2**(6): 662–667.

Sekaggya-Wiltshire, C, Braun, A von, Lamorde, M, Ledergerber, B, Buzibye, A, Henning, L, Musaazi, J, Gutteck, U, Denti, P, Kock, M de, Jetter, A, Byakika-Kibwika, P, Eberhard, N, Matovu, J, Joloba, M, Muller, D, Manabe, YC, Kamya, MR, Corti, N, Kambugu, A, Castelnuovo, B, Fehr, JS (2018) Delayed Sputum Culture Conversion in

Tuberculosis–Human Immunodeficiency Virus–Coinfected Patients With Low Isoniazid and Rifampicin Concentrations. *Clinical Infectious Diseases*, **67**(5): 708–716.

Sensi, P (1983) History of the Development of Rifampin. *Clinical Infectious Diseases*, **5**(Supplement_3): S402–S406.

Shah, NS, Auld, SC, Brust, JC, Mathema, B, Ismail, N, Moodley, P, Mlisana, K, Allana, S, Campbell, A, Mthiyane, T, Morris, N, Mpangase, P, Meulen, H van der, Omar, SV, Brown, TS, Narechania, A, Shaskina, E, Kapwata, T, Kreiswirth, B, Gandhi, NR (2017) Transmission of Extensively Drug-Resistant Tuberculosis in South Africa. *New England Journal of Medicine*, **376**(3): 243–253.

Shea, J, Halse, TA, Lapierre, P, Shudt, M, Kohlerschmidt, D, Van Roey, P, Limberger, R, Taylor, J, Escuyer, V, Musser, KA (2017) Comprehensive Whole-Genome Sequencing and Reporting of Drug Resistance Profiles on Clinical Cases of Mycobacterium tuberculosis in New York State. *Journal of Clinical Microbiology*, **55**(6): ed. by GA Land, 1871–1882.

Shen, G, Xue, Z, Shen, X, Sun, B, Gui, X, Shen, M, Mei, J, Gao, Q (2006) The study recurrent tuberculosis and exogenous reinfection, Shanghai, China. *Emerging infectious diseases*, **12**(11): 1776–8.

Sherman, DR, Mdluli, K, Hickey, MJ, Arain, TM, Morris, SL, Barry, CE, Stover, CK (1996) Compensatory ahpC Gene Expression in Isoniazid-Resistant Mycobacterium tuberculosis. *Science*, **272**(5268): 1641–1643.

Shimao, T (1987) Drug resistance in tuberculosis control. *Tubercle*, **68**(2 Suppl): 5–18.

Silva, PEA da, Groll, A von, Martin, A, Palomino, JC (2011) Efflux as a mechanism for drug resistance in Mycobacterium tuberculosis. *FEMS Immunology and Medical Microbiology*, **63**(1): 1–9.

Silva, PEA da, Machado, D, Ramos, D, Couto, I, Von Groll, A, Viveiros, M (2016) "Efflux Pumps in Mycobacteria: Antimicrobial Resistance, Physiological Functions, and Role in Pathogenicity". In: *Efflux-Mediated Antimicrobial Resistance in Bacteria*. Ed. by XZ Li, CA Elkins, HI Zgurskaya. Cham: Springer International Publishing, pp. 527–559.

Smith, T, Wolff, KA, Nguyen, L (2013) "Molecular biology of drug resistance in Mycobacterium tuberculosis". In: *Current Topics in Microbiology and Immunology*. Vol. 374, pp. 53–80.

Song, T, Park, Y, Shamputa, IC, Seo, S, Lee, SY, Jeon, HS, Choi, H, Lee, M, Glynne, RJ, Barnes, SW, Walker, JR, Batalov, S, Yusim, K, Feng, S, Tung, CS, Theiler, J, Via, LE, Boshoff, HIM, Murakami, KS, Korber, B, Barry, CE, Cho, SN (2014) Fitness costs of rifampicin resistance in Mycobacterium tuberculosis are amplified under conditions of

nutrient starvation and compensated by mutation in the β subunit of RNA polymerase. *Molecular Microbiology*, **91**(6): 1106–1119. eprint: NIHMS150003.

Soolingen, D van, Haas, PE de, Doorn, HR van, Kuijper, E, Rinder, H, Borgdorff, MW (2000) Mutations at amino acid position 315 of the katG gene are associated with high-level resistance to isoniazid, other drug resistance, and successful transmission of Mycobacterium tuberculosis in the Netherlands. *The Journal of infectious diseases*, **182**(6): 1788–90.

Spies, FS, Groll, A von, Ribeiro, AW, Ramos, DF, Ribeiro, MO, Dalla Costa, ER, Martin, A, Palomino, JC, Rossetti, ML, Zaha, A, Silva, PEA da (2013) Biological cost in Mycobacterium tuberculosis with mutations in the rpsL, rrs, rpoB, and katG genes. *Tuberculosis*, **93**(2): 150–154.

Springer, B, Lucke, K, Calligaris-Maibach, R, Ritter, C, Bottger, EC (2009) Quantitative Drug Susceptibility Testing of Mycobacterium tuberculosis by Use of MGIT 960 and EpiCenter Instrumentation. *Journal of Clinical Microbiology*, **47**(6): 1773–1780.

Sreevatsan, S, Pan, X, Stockbauer, KE, Connell, ND, Kreiswirth, BN, Whittam, TS, Musser, JM (1997) Restricted structural gene polymorphism in the Mycobacterium tuberculosis complex indicates evolutionarily recent global dissemination. *Proceedings of the National Academy of Sciences of the United States of America*, **94**(18): 9869–74.

Srivastava, S, Pasipanodya, JG, Meek, C, Leff, R, Gumbo, T (2011) Multidrug-resistant tuberculosis not due to noncompliance but to between-patient pharmacokinetic variability. *The Journal of infectious diseases*, **204**(12): 1951–9.

Starks, AM, Avilés, E, Cirillo, DM, Denkinger, CM, Dolinger, DL, Emerson, C, Gallarda, J, Hanna, D, Kim, PS, Liwski, R, Miotto, P, Schito, M, Zignol, M (2015) Collaborative Effort for a Centralized Worldwide Tuberculosis Relational Sequencing Data Platform. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, **61Suppl 3**(suppl 3): S141–6.

Steenwinkel, JE de, Kate, MT ten, Knegt, GJ de, Kremer, K, Aarnoutse, RE, Boeree, MJ, Verbrugh, HA, Soolingen, D van, Bakker-Woudenberg, IA (2012) Drug Susceptibility of Mycobacterium tuberculosis Beijing Genotype and Association with MDR TB. *Emerging Infectious Diseases*, **18**(4): 660–663.

Stefan, MA, Ugur, FS, Garcia, GA (2018) Source of the Fitness Defect in Rifamycin-Resistant Mycobacterium tuberculosis RNA Polymerase and the Mechanism of Compensation by Mutations in the β' Subunit. *Antimicrobial agents and chemotherapy*, **62**(6): e00164–18.

Streicher, EM, Bergval, I, Dheda, K, Böttger, EC, Gey van Pittius, NC, Bosman, M, Coetzee, G, Anthony, RM, Helden, PD van, Victor, TC, Warren, RM (2012) Mycobac-

terium tuberculosis Population Structure Determines the Outcome of Genetics-Based Second-Line Drug Resistance Testing. *Antimicrobial Agents and Chemotherapy*, **56**(5): 2420–2427.

Stucki, D, Ballif, M, Egger, M, Furrer, H, Altpeter, E, Battegay, M, Droz, S, Bruderer, T, Coscolla, M, Borrell, S, Zürcher, K, Janssens, JP, Calmy, A, Mazza Stalder, J, Jaton, K, Rieder, HL, Pfyffer, GE, Siegrist, HH, Hoffmann, M, Fehr, J, Dolina, M, Frei, R, Schrenzel, J, Böttger, EC, Gagneux, S, Fenner, L (2016) Standard Genotyping Overestimates Transmission of Mycobacterium tuberculosis among Immigrants in a Low-Incidence Country. *Journal of Clinical Microbiology*, **54**(7): ed. by KC Carroll, 1862–1870.

Stuckler, D, Basu, S, McKee, M, King, L (2008) Mass incarceration can explain population increases in TB and multidrug-resistant TB in European and central Asian countries. *Proceedings of the National Academy of Sciences*, **105**(36): 13280–13285.

Swaminathan, S, Ramachandran, G (2012) Role of pharmacogenomics in the treatment of tuberculosis: a review. *Pharmacogenomics and Personalized Medicine*, **5**(1): 89.

Szumowski, JD, Adams, KN, Edelstein, PH, Ramakrishnan, L (2012) "Antimicrobial Efflux Pumps and Mycobacterium tuberculosis Drug Tolerance: Evolutionary Considerations". In: *Current topics in microbiology and immunology*. Vol. 351. January, pp. 81–108.

Takiff, HE, Salazar, L, Guerrero, C, Philipp, W, Huang, WM, Kreiswirth, B, Cole, ST, Jacobs, WR, Telenti, A (1994) Cloning and nucleotide sequence of Mycobacterium tuberculosis gyrA and gyrB genes and detection of quinolone resistance mutations. *Antimicrobial Agents and Chemotherapy*, **38**(4): 773–780.

Telenti, A, Imboden, P, Marchesi, F, Lowrie, D, Cole, S, Colston, MJ, Matter, L, Schopfer, K, Bodmer, T (1993) Detection of rifampicin-resistance mutations in Mycobacterium tuberculosis. *Lancet (London, England)*, **341**(8846): 647–50.

The CRyPTIC Consortium and the 100000 Genomes Project (2018) Prediction of Susceptibility to First-Line Tuberculosis Drugs by DNA Sequencing. *New England Journal of Medicine*, NEJMoa1800474.

Turnbough, CL, Switzer, RL (2008) Regulation of Pyrimidine Biosynthetic Gene Expression in Bacteria: Repression without Repressors. *Microbiology and Molecular Biology Reviews*, **72**(2): 266–300.

Van Ingen, J, Aarnoutse, RE, Donald, PR, Diacon, AH, Dawson, R, Plemper Van Balen, G, Gillespie, SH, Boeree, MJ (2011) Why do we use 600 mg of rifampicin in tuberculosis treatment? *Clinical Infectious Diseases*, **52**(9):

Verbeke, LPC, Van den Eynden, J, Fierro, AC, Demeester, P, Fostier, J, Marchal, K (2015) Pathway Relevance Ranking for Tumor Samples through Network-Based Data Integration. *PLOS ONE*, **10**(7): ed. by S Patnaik, e0133503.

Verrall, AJ, G. Netea, M, Alisjahbana, B, Hill, PC, Crevel, R van (2014) Early clearance of Mycobacterium tuberculosis : a new frontier in prevention. *Immunology*, **141**(4): 506–513.

Vilchèze, C, Jacobs, WR (2014) Resistance to Isoniazid and Ethionamide in Mycobacterium tuberculosis: Genes, Mutations, and Causalities. *Microbiology spectrum*, **2**(4): MGM2–0014–2013.

Vos, M de, Muller, B, Borrell, S, Black, Pa, Helden, PD van, Warren, RM, Gagneux, S, Victor, TC (2013) Putative Compensatory Mutations in the rpoC Gene of Rifampin-Resistant Mycobacterium tuberculosis Are Associated with Ongoing Transmission. *Antimicrobial Agents and Chemotherapy*, **57**(2): 827–832.

Walker, TM, Ip, CL, Harrell, RH, Evans, JT, Kapatai, G, Dedicoat, MJ, Eyre, DW, Wilson, DJ, Hawkey, PM, Crook, DW, Parkhill, J, Harris, D, Walker, AS, Bowden, R, Monk, P, Smith, EG, Peto, TE (2013) Whole-genome sequencing to delineate Mycobacterium tuberculosis outbreaks: a retrospective observational study. *The Lancet Infectious Diseases*, **13**(2): 137–146.

Walker, TM, Kohl, TA, Omar, SV, Hedge, J, Del Ojo Elias, C, Bradley, P, Iqbal, Z, Feuerriegel, S, Niehaus, KE, Wilson, DJ, Clifton, DA, Kapatai, G, Ip, CLC, Bowden, R, Drobniewski, FA, Allix-Béguec, C, Gaudin, C, Parkhill, J, Diel, R, Supply, P, Crook, DW, Smith, EG, Walker, AS, Ismail, N, Niemann, S, Peto, TEA (2015) Whole-genome sequencing for prediction of Mycobacterium tuberculosis drug susceptibility and resistance: a retrospective cohort study. *The Lancet Infectious Diseases*, **15**(10): 1193–1202.

Wang, F, Cassidy, C, Sacchettini, JC (2006) Crystal Structure and Activity Studies of the Mycobacterium tuberculosis -Lactamase Reveal Its Critical Role in Resistance to -Lactam Antibiotics. *Antimicrobial Agents and Chemotherapy*, **50**(8): 2762–2771.

Warner, DF, Mizrahi, V (2006) Tuberculosis chemotherapy: The influence of bacillary stress and damage response pathways on drug efficacy. *Clinical Microbiology Reviews*, **19**(3): 558–570.

Warren, JL, Grandjean, L, Moore, DAJ, Lithgow, A, Coronel, J, Sheen, P, Zelner, JL, Andrews, JR, Cohen, T (2018) Investigating spillover of multidrug-resistant tuberculosis from a prison: a spatial and molecular epidemiological analysis. *BMC Medicine*, **16**(1): 122.

Warrier, T, Kapilashrami, K, Argyrou, A, Ioerger, TR, Little, D, Murphy, KC, Nandakumar, M, Park, S, Gold, B, Mi, J, Zhang, T, Meiler, E, Rees, M, Somersan-Karakaya,

S, Porras-De Francisco, E, Martinez-Hoyos, M, Burns-Huang, K, Roberts, J, Ling, Y, Rhee, KY, Mendoza-Losana, A, Luo, M, Nathan, CF (2016) N-methylation of a bactericidal compound as a resistance mechanism in Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences of the United States of America*, **113**(31): E4523–30.

Warrier, T, Martinez-Hoyos, M, Marin-Amieva, M, Colmenarejo, G, Porras-De Francisco, E, Alvarez-Pedraglio, AI, Fraile-Gabaldon, MT, Torres-Gomez, PA, Lopez-Quezada, L, Gold, B, Roberts, J, Ling, Y, Somersan-Karakaya, S, Little, D, Cammack, N, Nathan, C, Mendoza-Losana, A (2015) Identification of Novel Anti-mycobacterial Compounds by Screening a Pharmaceutical Small-Molecule Library against Nonreplicating Mycobacterium tuberculosis. *ACS Infectious Diseases*, **1**(12): 580–585.

Welin, A, Raffetseder, J, Eklund, D, Stendahl, O, Lerm, M (2011) Importance of phagosomal functionality for growth restriction of Mycobacterium tuberculosis in primary human macrophages. *Journal of innate immunity*, **3**(5): 508–18.

Werngren, J, Hoffner, SE (2003) Drug-Susceptible Mycobacterium tuberculosis Beijing Genotype Does Not Develop Mutation-Conferred Resistance to Rifampin at an Elevated Rate. *Journal of Clinical Microbiology*, **41**(4): 1520–1524.

WHO (2013) *Automated Real-Time Nucleic Acid Amplification Technology for Rapid and Simultaneous Detection of Tuberculosis and Rifampicin Resistance: Xpert MTB/RIF Assay for the Diagnosis of Pulmonary and Extrapulmonary TB in Adults and Children: Policy Update*, 1–79.

WHO (2014) Drug Resistance TB Surveillance and Response – Supplement Global Tuberculosis Report 2014, 32.

Wilson, LG (1990) The historical decline of tuberculosis in Europe and America: Its causes and significance. *Journal of the History of Medicine and Allied Sciences*, **45**(3): 366–396.

World Health Organization (1991) *Guidelines for TB treatment in adults and children in national TB programmes*. Geneva.

World Health Organization (1997) *TB : WHO report on the tuberculosis epidemic 1997*. Geneva.

World Health Organization (2010) *Treatment of tuberculosis: Guidelines*. 4th Editio, 1–160.

World Health Organization (2014a) *Companion handbook to the WHO guidelines for the programmatic management of drug-resistant tuberculosis*. Geneva, 464.

World Health Organization (2014b) *Extensive review of tuberculosis prevention, control and care in Georgia 6–14 November 2014*. Tech. rep. Copenhagen.

World Health Organization (2015) *Analysis of the epidemiological impact of tuberculosis in Georgia*. Tech. rep. Copenhagen.

World Health Organization (2016a) *Global tuberculosis report 2016*. Geneva, 211.

World Health Organization (2016b) *WHO Guideline: The use of molecular line probe assays for the detection of resistance to isoniazid and rifampicin*. Geneva, 56.

World Health Organization (2018a) *Global Tuberculosis Report 2018*. Geneva: World Health Organization.

World Health Organization (2018b) *Technical Report on critical concentrations for drug susceptibility testing of medicines used in the treatment of drug-resistant tuberculosis*, 1–106.

World Health Organization (2018c) *The use of next-generation sequencing technologies for the detection of mutations associated with drug resistance in Mycobacterium tuberculosis complex: technical guide*. Geneva, 128.

World Health Organization, European Centre for Disease Prevention and Control (2018) *European Centre for Disease Prevention and Control/WHO Regional Office for Europe. Tuberculosis surveillance and monitoring in Europe 2018 – 2016 data.* 1st. Stockholm, 129.

Xu, C, Kreiswirth, BN, Sreevatsan, S, Musser, JM, Drlica, K (1996) Fluoroquinolone Resistance Associated with Specific Gyrase Mutations in Clinical Isolates of Multidrug-Resistant Mycobacterium tuberculosis. *Journal of Infectious Diseases*, **174**(5): 1127–1130.

Yang, C, Luo, T, Shen, X, Wu, J, Gan, M, Xu, P, Wu, Z, Lin, S, Tian, J, Liu, Q, Yuan, Z, Mei, J, DeRiemer, K, Gao, Q (2016) Transmission of multidrug-resistant Mycobacterium tuberculosis in Shanghai, China: a retrospective observational study using whole-genome sequencing and epidemiological investigation. *The Lancet Infectious Diseases*, **3099**(16): 1–10.

Yates, TA, Khan, PY, Knight, GM, Taylor, JG, McHugh, TD, Lipman, M, White, RG, Cohen, T, Cobelens, FG, Wood, R, Moore, DAJ, Abubakar, I (2016) The transmission of Mycobacterium tuberculosis in high burden settings. *The Lancet Infectious Diseases*, **16**(2): 227–238.

Yew, W, Wong, C, Lee, J, Wong, P, Chan, C (1995) Do β-lactam-β-lactamase inhibitor combinations have a place in the treatment of multidrug-resistant pulmonary tuberculosis? *Tubercle and Lung Disease*, **76**(1): 90–92.

Yu, S, Girotto, S, Lee, C, Magliozzo, RS (2003) Reduced Affinity for Isoniazid in the S315T Mutant of Mycobacterium tuberculosis KatG Is a Key Factor in Antibiotic Resistance. *Journal of Biological Chemistry*, **278**(17): 14769–14775.

Yu, X, Jiang, G, Li, H, Zhao, Y, Zhang, H, Zhao, L, Ma, Y, Coulter, C, Huang, H (2011) Rifampin Stability in 7H9 Broth and Lowenstein-Jensen Medium. *Journal of Clinical Microbiology*, **49**(3): 784–789.

Zaunbrecher, MA, Sikes, RD, Metchock, B, Shinnick, TM, Posey, JE (2009) Overexpression of the chromosomally encoded aminoglycoside acetyltransferase eis confers kanamycin resistance in Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences*, **106**(47): 20004–20009.

Zhang, H, Li, D, Zhao, L, Fleming, J, Lin, N, Wang, T, Liu, Z, Li, C, Galwey, N, Deng, J, Zhou, Y, Zhu, Y, Gao, Y, Wang, T, Wang, S, Huang, Y, Wang, M, Zhong, Q, Zhou, L, Chen, T, Zhou, J, Yang, R, Zhu, G, Hang, H, Zhang, J, Li, F, Wan, K, Wang, J, Zhang, Xe, Bi, L (2013) Genome sequencing of 161 Mycobacterium tuberculosis isolates from China identifies genes and intergenic regions associated with drug resistance. *Nature genetics*, **45**(10): 1255–60.

Zhang, X, Yin, J, Li, H, Li, S, Walley, J, Zou, G, Zhang, Z, Wei, X (2015a) Diagnostic and treatment delays of multidrug-resistant tuberculosis before initiating treatment: a cross-sectional study. *Tropical Medicine & International Health*, **20**(11): 1431–1437.

Zhang, Y, Yew, WW (2015b) Mechanisms of drug resistance in Mycobacterium tuberculosis: update 2015. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease*, **19**(11): 1276–89.

Zhao, F, Wang, XD, Erber, LN, Luo, M, Guo, AZ, Yang, SS, Gu, J, Turman, BJ, Gao, YR, Li, DF, Cui, ZQ, Zhang, ZP, Bi, LJ, Baughn, AD, Zhang, XE, Deng, JY (2014) Binding pocket alterations in dihydrofolate synthase confer resistance to para-aminosalicylic acid in clinical isolates of Mycobacterium tuberculosis. *Antimicrobial Agents and Chemotherapy*, **58**(3): 1479–1487.

Zumla, AI, Gillespie, SH, Hoelscher, M, Philips, PPJ, Cole, ST, Abubakar, I, McHugh, TD, Schito, M, Maeurer, M, Nunn, AJ (2014) New antituberculosis drugs, regimens, and adjunct therapies: Needs, advances, and future prospects. *The Lancet Infectious Diseases*, **14**(4): 327–340.

Zumla, A, Chakaya, J, Hoelscher, M, Ntoumi, F, Rustomjee, R, Vilaplana, C, Yeboah-Manu, D, Rasolof, V, Munderi, P, Singh, N, Aklillu, E, Padayatchi, N, Macete, E, Kapata, N, Mulenga, M, Kibiki, G, Mfinanga, S, Nyirenda, T, Maboko, L, Garcia-Basteiro, A, Rakotosamimanana, N, Bates, M, Mwaba, P, Reither, K, Gagneux, S, Edwards, S, Mfinanga, E, Abdulla, S, Cardona, PJ, Russell, JB, Gant, V, Noursadeghi, M, Elkington, P, Bonnet, M, Menendez, C, Dieye, TN, Diarra, B, Maiga, A, Aseffa, A, Parida, S, Wejse, C, Petersen, E, Kaleebu, P, Oliver, M, Craig, G, Corrah, T, Tientcheu,

L, Antonio, M, Rao, M, McHugh, TD, Sheikh, A, Ippolito, G, Ramjee, G, Kaufmann, SHE, Churchyard, G, Steyn, A, Grobusch, M, Sanne, I, Martinson, N, Madansein, R, Wilkinson, RJ, Mayosi, B, Schito, M, Wallis, RS, Maeurer, M (2015) Towards host-directed therapies for tuberculosis. *Nature Reviews Drug Discovery*, **14**(8): 511–512.

Zürcher, K, Ballif, M, Zwahlen, M, Rieder, HL, Egger, M, Fenner, L (2016) Tuberculosis Mortality and Living Conditions in Bern, Switzerland, 1856-1950. *PLOS ONE*, **11**(2): ed. by DW Dowdy, e0149195.

List of Figures

List of Tables

A. Appendix

A1. Supplement to Chapter 4

A1.1. Materials and Methods

Culture conditions

Strains were isolated from clinical material following established decontamination procedures. Bacteria were cultured in BD MGIT 960 tubes and Middlebrook 7H10 agar media. Strains collected as a part of the IeDEA study were cultured on Löwenstein-Jensen slants and sent to the National Center for Mycobacteria (University of Zürich, Zürich, Switzerland) for phenotypic drug susceptibility testing.

Phenotypic drug susceptibility testing – BD MGIT 960

DST for first- and second-line anti-TB drugs was performed using the Becton Dickinson (BD) MGIT 960 and EpiCenter devices equipped with the TB eXiST module (BD, Franklin Lakes, NJ, USA) (Springer *et al.*, 2009). The MGIT 960 system was used for primary bacterial isolation and DST for first-line drugs (rifampicin, isoniazid, ethambutol, pyrazinamide) as recommended by the manufacturer, using currently established epidemiological cut-offs (ECOFF – epidemiological cut-off, i.e. the highest observed wild-type MIC (Ängeby *et al.*, 2012)) summarized in Table 4.1 in the main text. MGIT tubes enriched with 0.8 ml of supplement (MGIT 960 OADC supplement; BD) were inoculated with 0.1 ml of the drug solution (Table A.1.2) and 0.5 ml of the test strain suspension. For the drug-free growth control tube, the *M. tuberculosis* suspension was diluted 1:100 in sterile saline and 0.5 ml was inoculated into the tube (proportion testing).

Phenotypic drug susceptibility testing - Proportion method by 7H10 agar dilution

M. tuberculosis strains were grown in Middlebrook 7H9 liquid broth with 10 % OADC supplement (BD) until McFarland 0.5 was reached. Using a replicator micropipettor robot VIAFLOW96 (Integra Biosciences, Zizers, Switzerland), 1 μl of the culture suspension was inoculated onto Middlebrook 7H10 agar plates with twofold serially diluted drug concentrations to determine the MIC (Table A.1.2), whereby the MIC is defined as the lowest drug concentration that inhibits growth of more than 99 % the bacteria during 21 days of incubation at 37° C (Sirgel *et al.*, 2009). Agar plates were read automatically using an AID Microplate reader and automated software developed by AID (AID Diagnostika, Strassberg, Germany). Phenotypic DST for all strains was performed at the Swiss National Center for Mycobacteria (University of Z[ü]rich, Z[ü]rich, Switzerland).

Whole genome sequencing

Libraries were prepared using the Illumina Nextera XT kit and sequenced on an Illumina HiSeq 2500, generating 125 bp paired-end reads. Sequencing was performed at the genomics facility of the ETHZ/University of Basel in Basel, Switzerland and at the Broad Institute, Cambridge, Massachusetts, United States. The raw data was processed with an in-house python pipeline as follows: reads were adaptor clipped and quality trimmed with Trimmomatic (v.0.33), whereby resulting reads < 20 bp were discarded. Overlapping paired-end reads were merged using SeqPrep (https://github.com/jstjohn/SeqPrep). The processed reads were subsequently mapped to a reconstructed hypothetical MTBC ancestor (Comas *et al.*, 2010) with BWA (v.0.7.12) (Li *et al.*, 2009a). Duplicated reads were marked with Picard (v.2.1.1) (https://github.com/broadinstitute/picard) using the MarkDuplicates module. Local realignments of reads around indels were performed with the GATK (v.3.4.0) modules RealignerTargetCreator and IndelRealigner (McKenna *et al.*, 2010). Pileups were generated with Samtools (v.1.2) (Li *et al.*, 2009b) and SNVs were subsequently called with VarScan (v.2.4.1) (Koboldt *et al.*, 2012) using the following thresholds: minimum mapping & minimum base quality of 20, minimum read depth of 7x at a given position. For a SNV to be called, the alternative base call needed to be supported by at lest 5 reads without strand bias. Furthermore, SNVs were considered as fixed when they reached a frequency of ≥ 90 %. The position was called as ancestral when the frequency was found to be < 10 %. SNVs were annotated with SnpEff (v.4.11) (Cingolani *et al.*, 2012) corresponding to the *M. tuberculosis* H37Rv reference annotation

(NC_000962.3). SNVs in regions that share a minimum of 50 bp of sequence identity with other regions in the genome were excluded (Stucki, 2015).

Variable SNV alignment and phylogenetic analysis

A variable SNV alignment was generated by concatenating all filtered SNVs in the dataset, whereby the IUPAC nucleotide ambiguity codes were used for unfixed positions (10 % ≤ variant frequency ≤ 90 %). Positions were considered variable if at least one genome had SNV-call at the position in question. If the SNV fell into an excluded region (see above) or was covered by less than 7 reads it was encoded as X in the alignment. If there was no sequence information at all available for a position, it was encoded as a gap. Furthermore, positions known to be involved in drug resistance were not considered in the alignment. The variable SNV alignment was used to infer a maximum likelihood phylogeny with RAxML (v.8.2.8) (Stamatakis, 2014), using the general time-reversible model of sequence evolution and branch support values were inferred by bootstrapping the highest scoring maximum likelihood tree (1000 pseudoreplicates). *Mycobacterium canettii* (SRR011186) was used to root the phylogeny. The genomes were classified into main and sublineages based on the presence of previously established markers (Coll *et al.*, 2014). Average genetic distances between strains based on fixed single nucleotide variants (SNV) were calculated with the ape package (v.4.1) (Paradis *et al.*, 2004) for R (v.3.3.3) (R Core Team, 2013). Two strains were defined as clustered if their average genetic distance was ≤ 12 SNVs (Walker *et al.*, 2013) and one strain was subsequently omitted from the analysis at random. Phylogenetic trees with associated metadata were visualized with the R package ggtree (v.1.6.11) (Yu *et al.*, 2017).

WGS-based resistance profile inference

The WGS data was screened for non-synonymous mutations and indels in genes known to be involved in drug resistance for the drugs assayed in *M. tuberculosis* (Gygli *et al.*, 2017; Safi *et al.*, 2013). Phylogenetic markers in drug resistance-associated genes (Table 4.2 – main text) shared by all strains belonging to a main- or sublineage (Coll *et al.*, 2014; Stucki *et al.*, 2016) (Table A.1.1), as well as previously described phylogenetic markers (Coll *et al.*, 2015), including main lineage markers derived from an unpublished collection of 400 phylogenetically diverse strains, were removed. After filtering for phylogenetic markers, all fixed non-synonymous polymorphisms in resistance-related genes were treated as linked to drug resistance. Where necessary, WGS results were complemented by molecular DST approaches as described previously (Ritter *et al.*, 2014).

Data availability

The datasets generated and analysed in the study were deposited at the National Center National Center for Biotechnology Information (NCBI) under the BioProject IDs PRJNA454477 (http://www.ncbi.nlm.nih.gov/bioproject/454477) and PRJNA300846 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA300846). The quantitative readout of the 7H10 agar dilution-based DST and the results of the MGIT 960-based testing at the WHO-defined critical concentration are available upon request.

A1.2. Supplementary Tables

Table A1.1.: **List of major phylogenetic markers in drug-resistance related genes.**

Gene	Rv-Number	Position	Ref. nt.	Alt nt.	Codon change	Lineage	Coll sublineage	Comments
embC	Rv3793	4242182	G	T	Ala774Ser	L4	L4.3.3	Sublineage of L4.3.3
gid	Rv3919c	4407967	T	C	Leu79Ser	L4	L4.1	Sublineage of L4.1

Table A1.2.: **List of drug concentrations tested for MGIT 960 & 7H10 agar dilution.** ND = not determined.

Antibiotic	MGIT 960 concentrations (mg/L) Training set (n = 56)	MGIT 960 concentrations (mg/L) Test set (n = 120)	7H10 agar dilution concentration ranges (mg/L), $\log_2$ steps, Training set (n = 56)	7H10 agar dilution concentration, ranges (mg/L), $\log_2$ steps, Test set (n = 120)
Ethionamide	1.25, 2.5, 5, 10, 25	ND	0.25-128	0.25-256
Ethambutol	1.25, 2.5, 5, 12.5, 50	5	0.5-64	0.5-64
Capreomycin	2.5, 5, 25	ND	0.5-64	0.5-64
Streptomycin	1, 4, 20	ND	0.06-128	0.06-128
Kanamycin A	1, 2, 4, 20	ND	0.06-128	0.06-128
Amikacin	1, 4, 20	1	0.06 - 128	0.06-128
Moxifloxacin	0.25, 0.5, 2.5, 7.5	0.25	0.016-32	0.016-32
Isoniazid	0.1, 1, 3, 10	0.1, 1	0.004-32	0.004-32
Rifampicin	1, 4, 20	1	0.016-256	0.016-256
Rifabutin	0.1, 0.4, 2	ND	0.004-32	0.004-32
Pyrazinamide	100	100	ND	ND

A1.3. Supplementary Figures

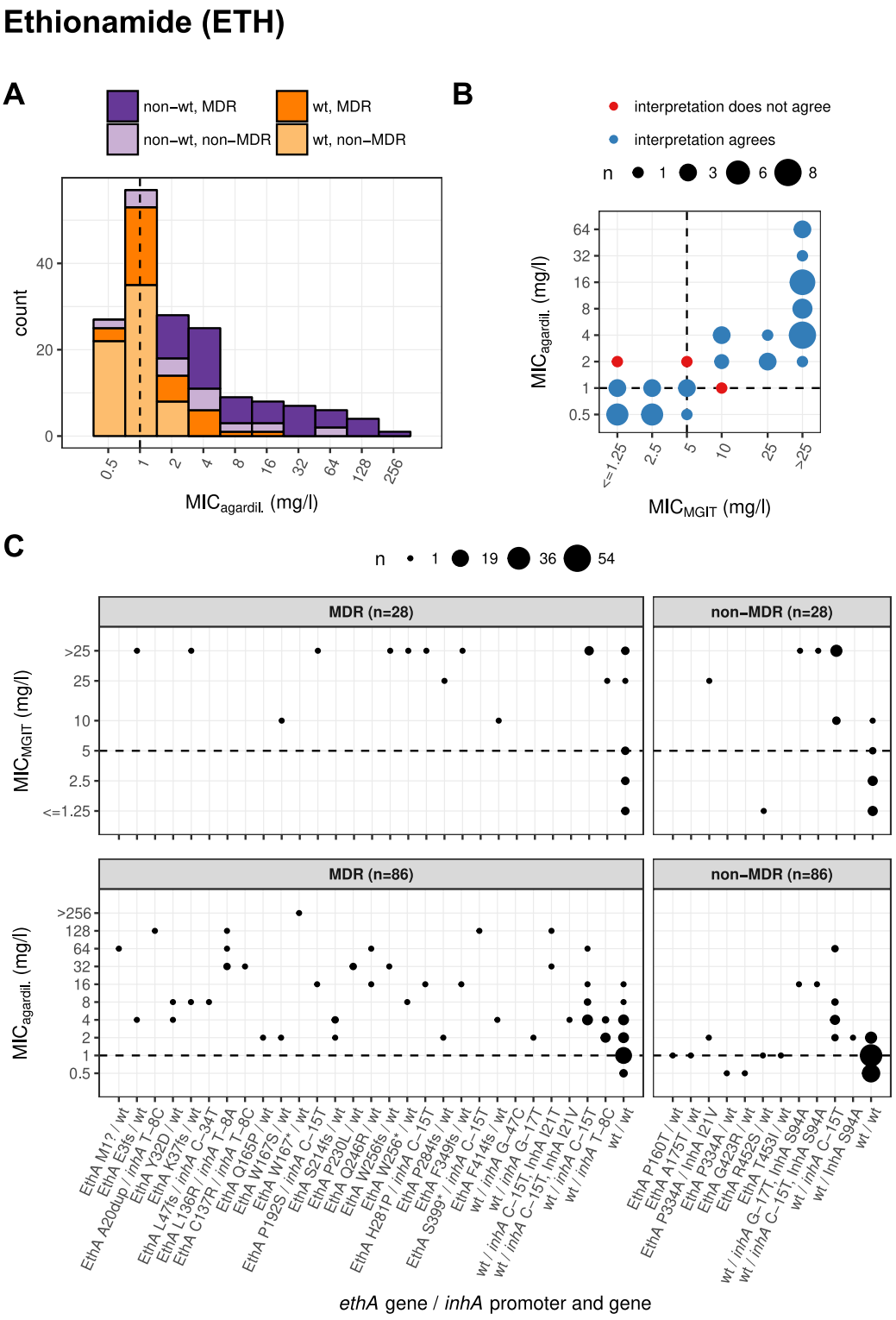

Figure A1.1.: **Summary of DST results for ethionamideA** Histogram of 7H10 agar dilution MICs for **ethionamide**. **B** Method agreement between phenotypic DST by MGIT 960 and 7H10 agar dilution. **C** MICs of *M. tuberculosis* strains harboring resistance mutations or for wt. Top panel MGIT 960, bottom panel 7H10 agar dil.. "M1?" indicates an initiator codon variant, "fs" indicates a frame shift variant, "*" indicates a stop-gain variant. Dashed lines indicate the ECOFF.

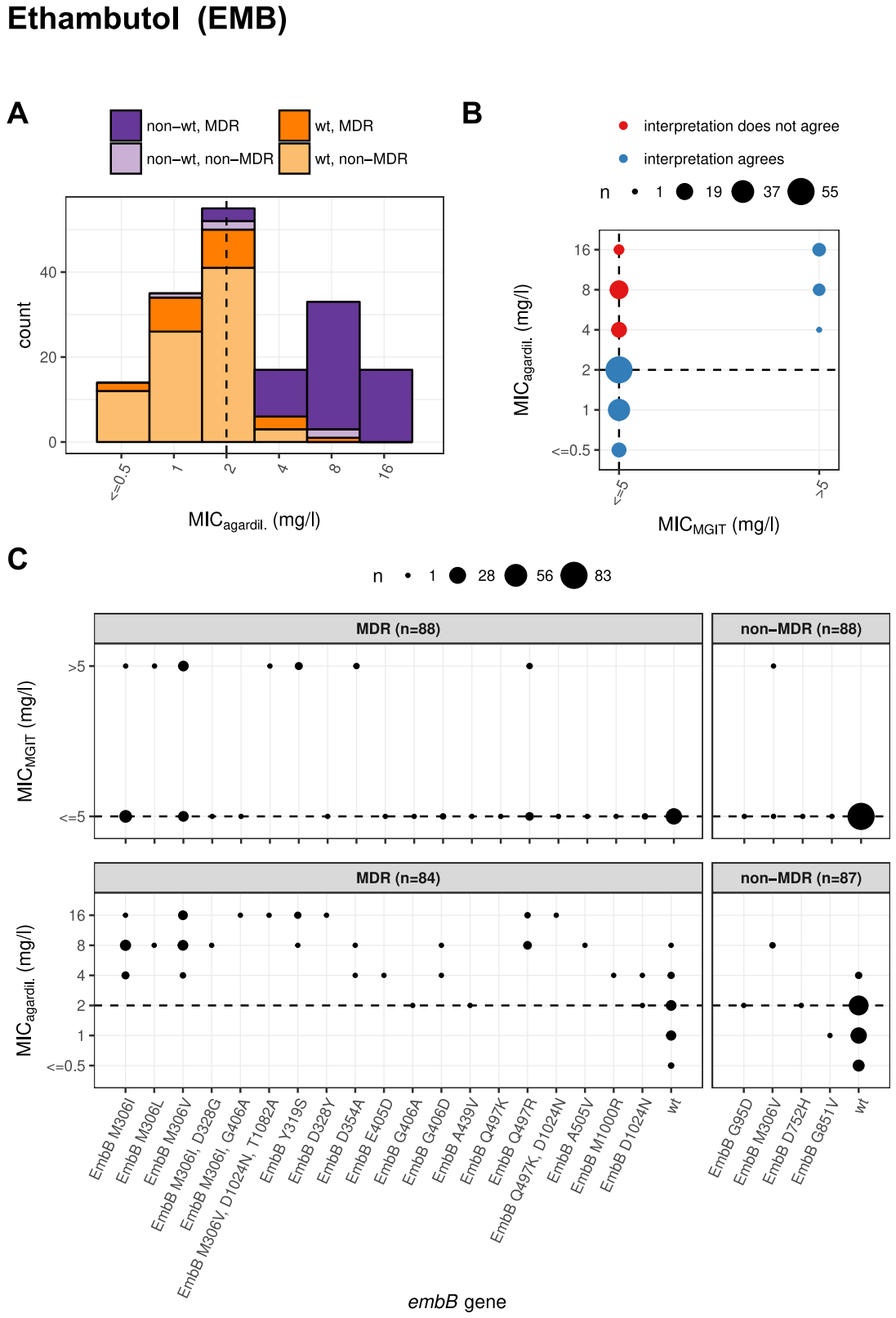

Figure A1.2.: **Summary of DST results for ethambutolA** Histogram of 7H10 agar dilution MICs for ethambutol resistance. **B** method agreement between phenotypic DST by MGIT 960 and 7H10 agar dilution. **C** MICs of *M. tuberculosis* strains harboring resistance mutations in target genes or for wt. Top panel MGIT 960, bottom panel 7H10 agar dilution. In all panels, dashed lines indicate the ECOFF.

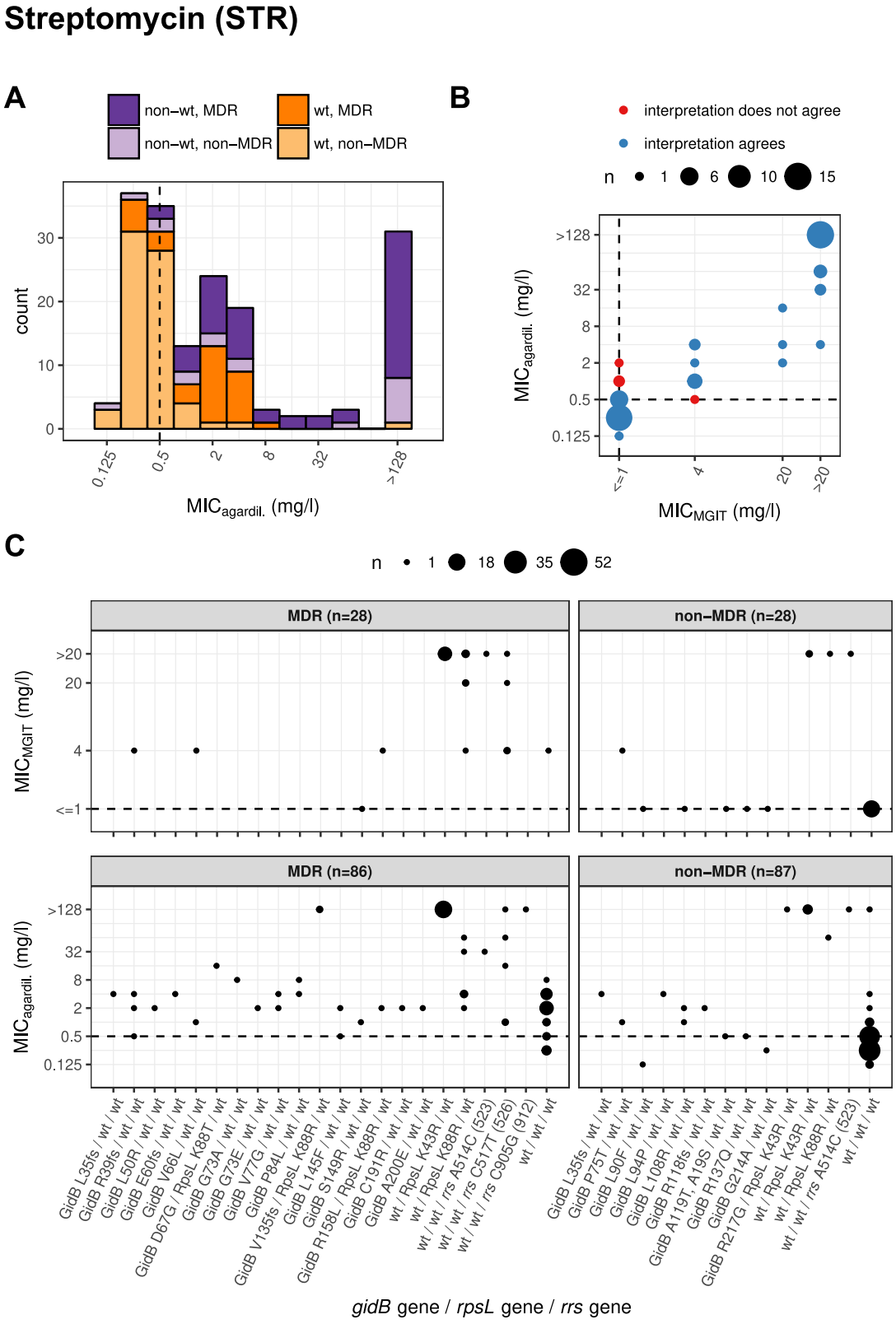

Figure A1.3.: **Summary of DST results for streptomycin A** Histogram of 7H10 agar dilution MICs for **streptomycin** resistance. **B** method agreement between phenotypic DST by MGIT 960 and 7H10 agar dilution. **C** MICs of *M. tuberculosis* strains harboring resistance mutations in target genes or for wt. Numbers in parentheses indicate the corresponding *Escherichia coli* nucleotide numbering. Top panel MGIT 960, bottom panel 7H10 agar dilution. "fs" indicates a frame shift variant. In all panels, dashed lines indicate the ECOFF.

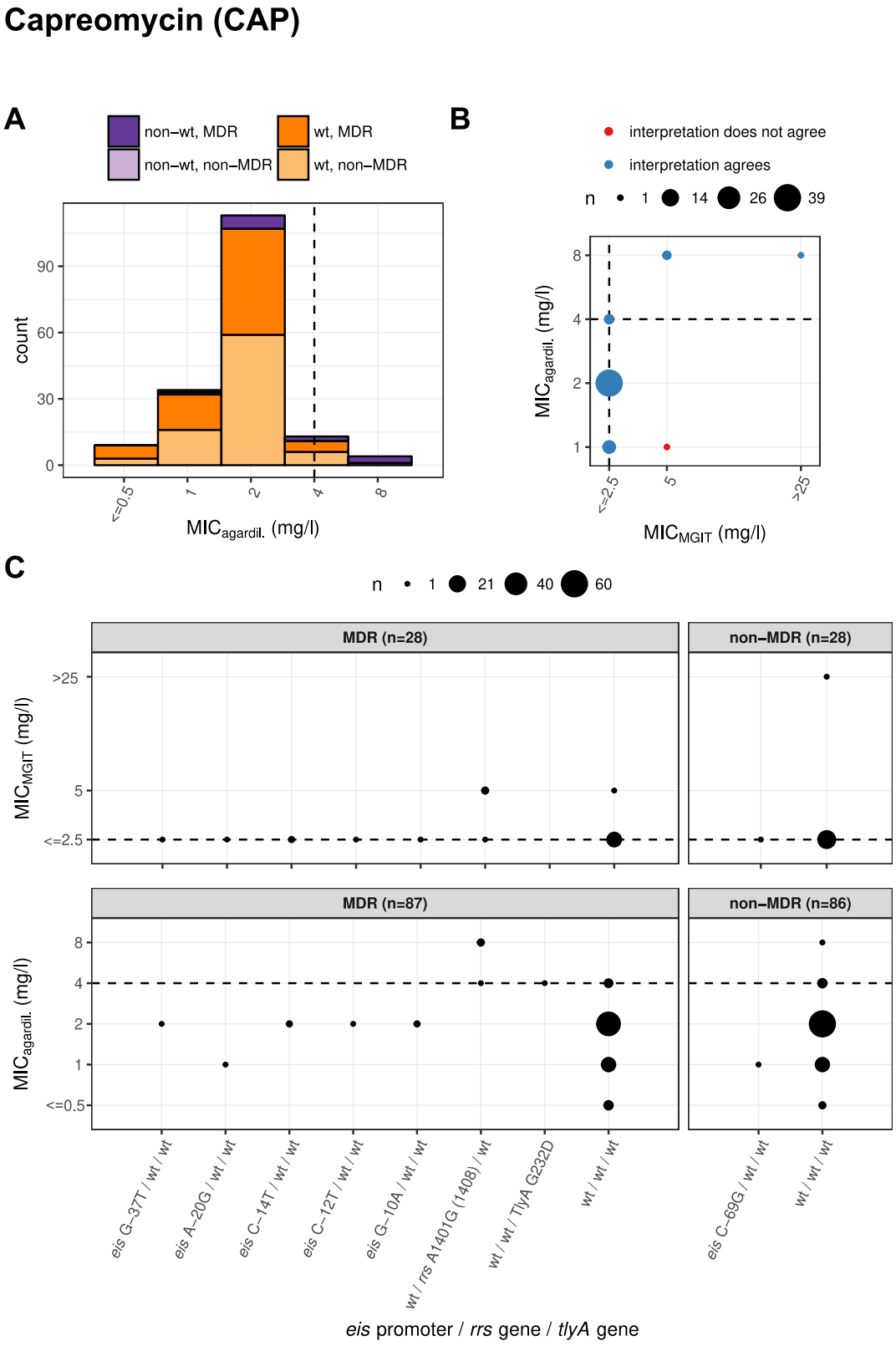

Figure A1.4.: **Summary of DST results for capreomycin. A** Histogram of 7H10 agar dilution MICs for capreomycin resistance. **B** method agreement between phenotypic DST by MGIT 960 and 7H10 agar dilution. **C** MICs of *M. tuberculosis* strains harboring resistance mutations in target genes or for wt. Numbers in parentheses indicate the corresponding *Escherichia coli* nucleotide numbering. Top panel MGIT 960, bottom panel 7H10 agar dilution. In all panels, dashed lines indicate the ECOFF.

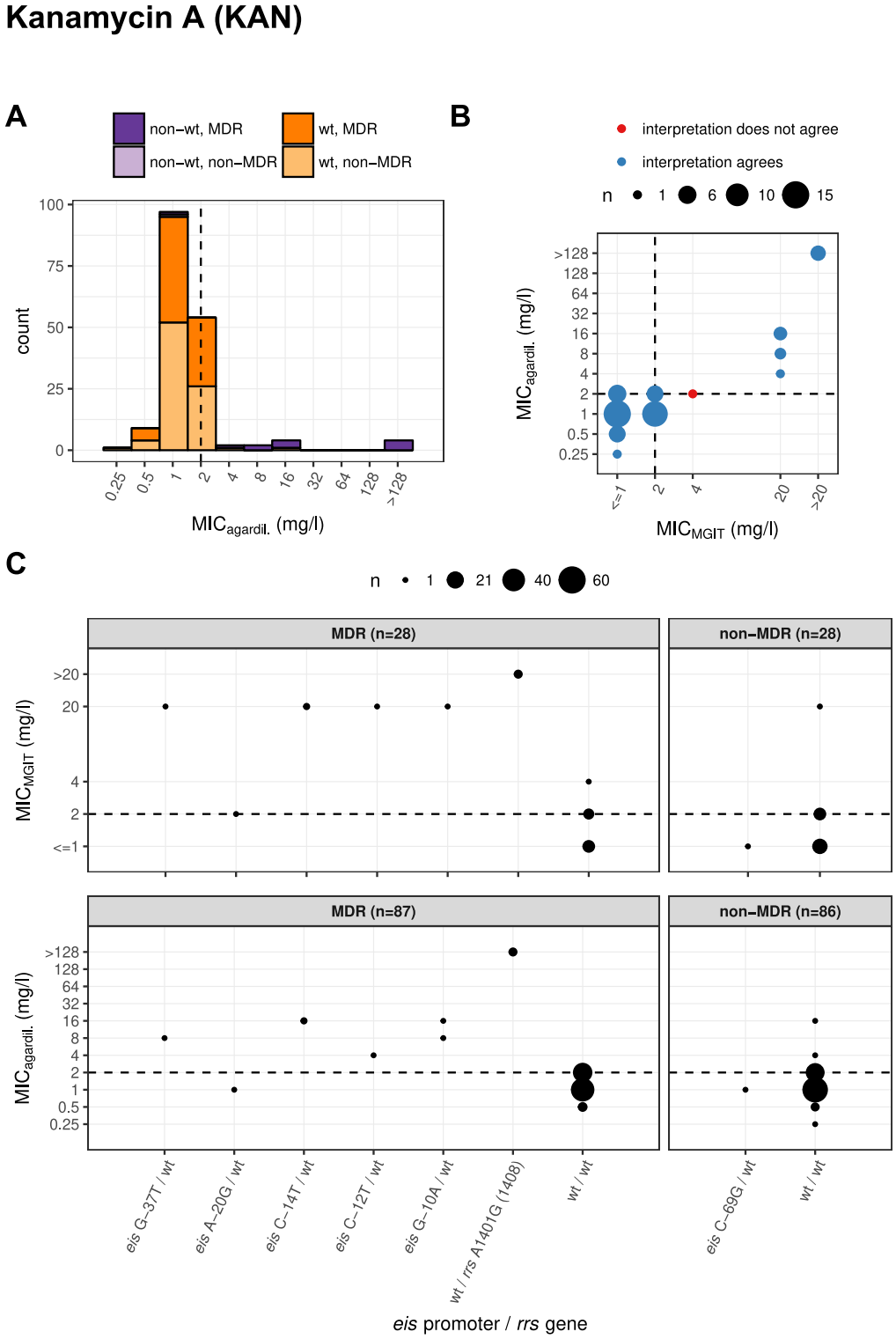

Figure A1.5.: **Summary of DST results for kanamycin A. A** Histogram of 7H10 agar dilution MICs for kanamycin A resistance. **B** method agreement between phenotypic DST by MGIT 960 and 7H10 agar dilution. **C** MICs of *M. tuberculosis* strains harboring resistance mutations in target genes or for wt.Numbers in parentheses indicate the corresponding *Escherichia coli* nucleotide numbering. Top panel MGIT 960, bottom panel 7H10 agar dilution. Numbers in parentheses indicate the corresponding *Escherichia coli* nucleotide number. In all panels, dashed lines indicate the ECOFF.

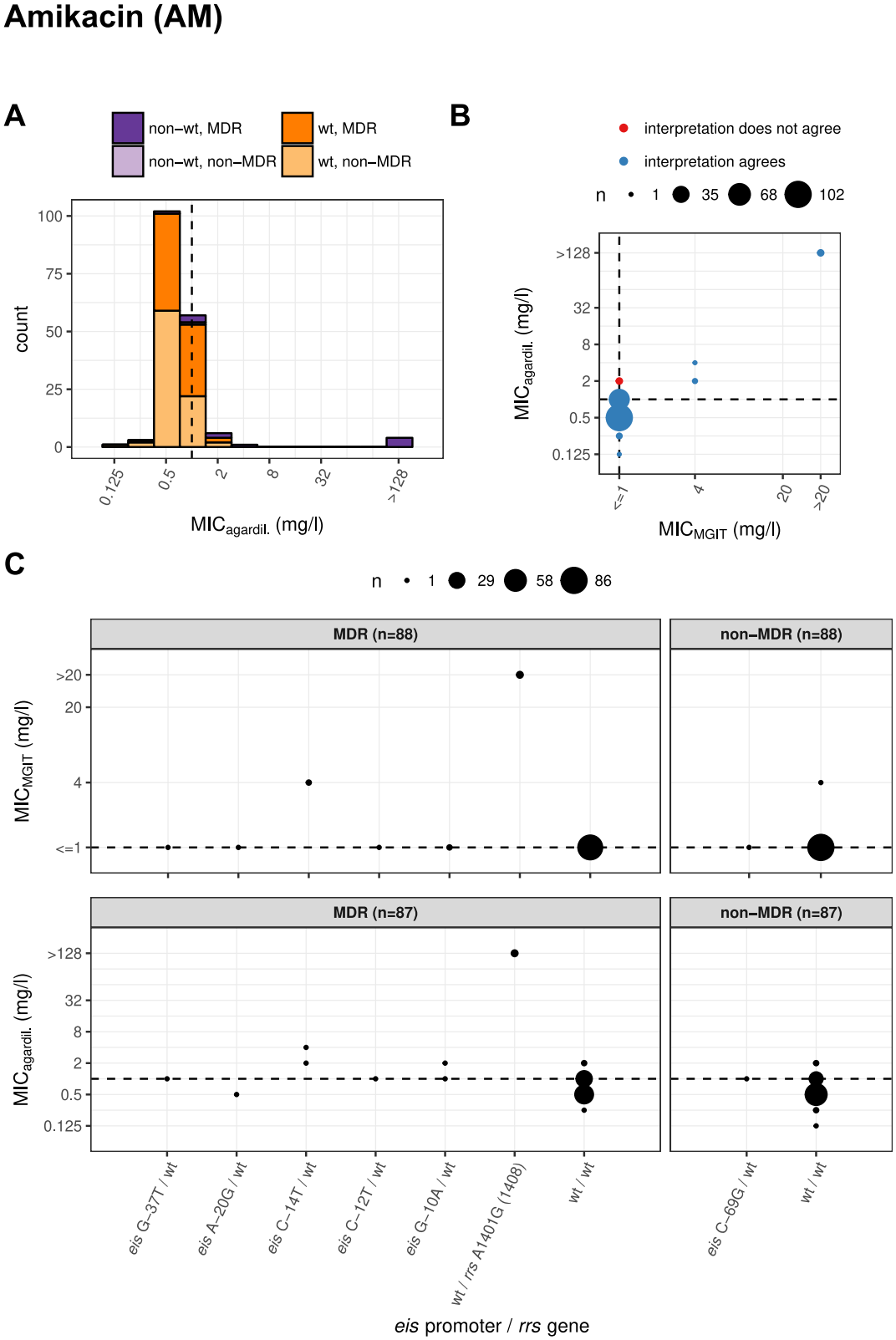

Figure A1.6.: **Summary of DST results for amikacin.** **A** Histogram of 7H10 agar dilution MICs for amikacin resistance. **B** method agreement between phenotypic DST by MGIT 960 and 7H10 agar dilution. **C** MICs of M. tuberculosis strains harboring resistance mutations in target genes or for wt. Numbers in parentheses indicate the corresponding *Escherichia coli* nucleotide numbering. Top panel MGIT 960, bottom panel 7H10 agar dilution. In all panels, dashed lines indicate the ECOFF.

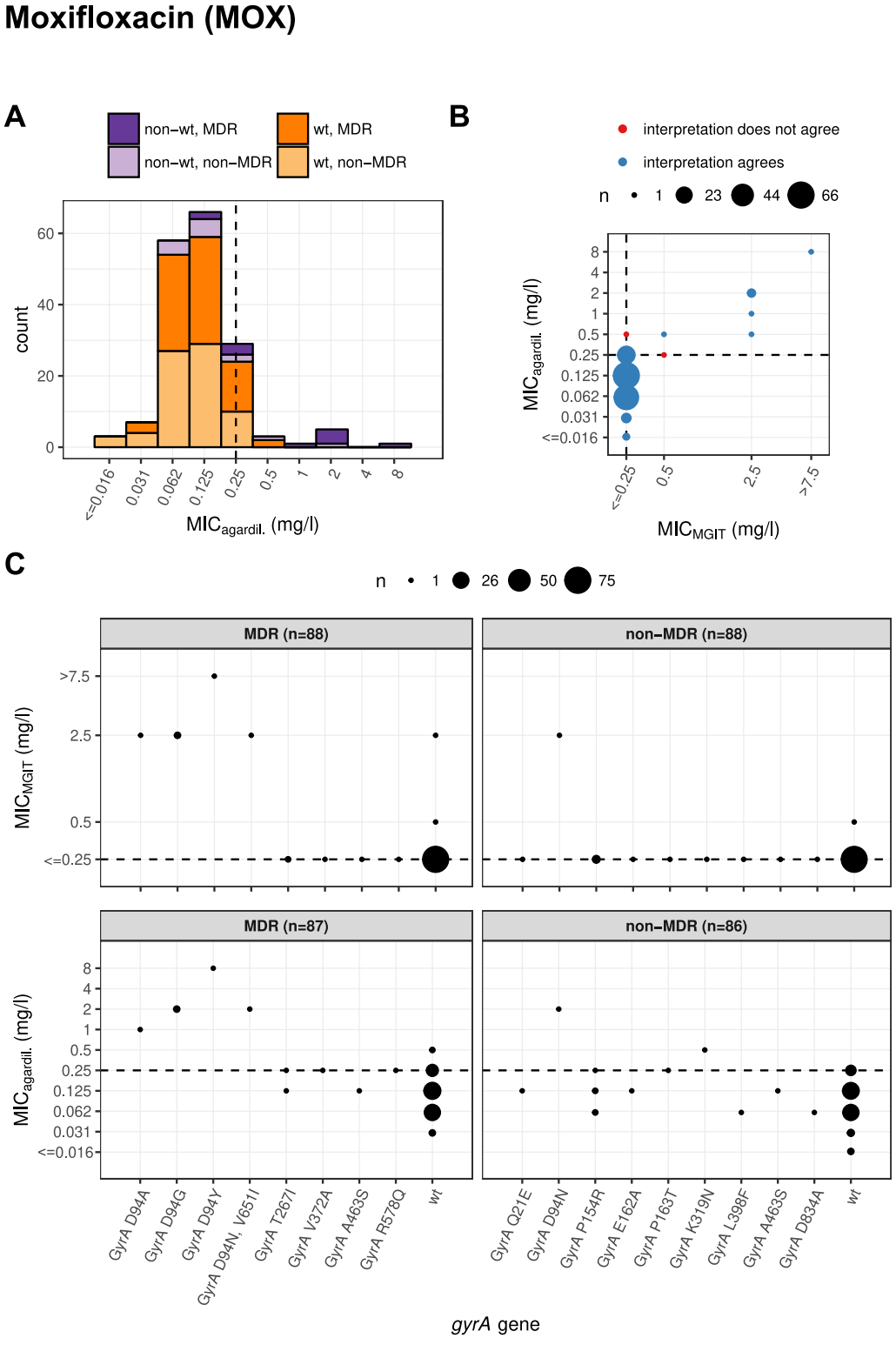

Figure A1.7.: **Summary of DST results for moxifloxacin. A** Histogram of 7H10 agar dilution MICs for moxifloxacin resistance. **B** method agreement between phenotypic DST by MGIT 960 and 7H10 agar dilution. **C** MICs of *M. tuberculosis* strains harboring resistance mutations in target genes or for wt. Top panel MGIT 960, bottom panel 7H10 agar dilution. In all panels, dashed lines indicate the ECOFF.

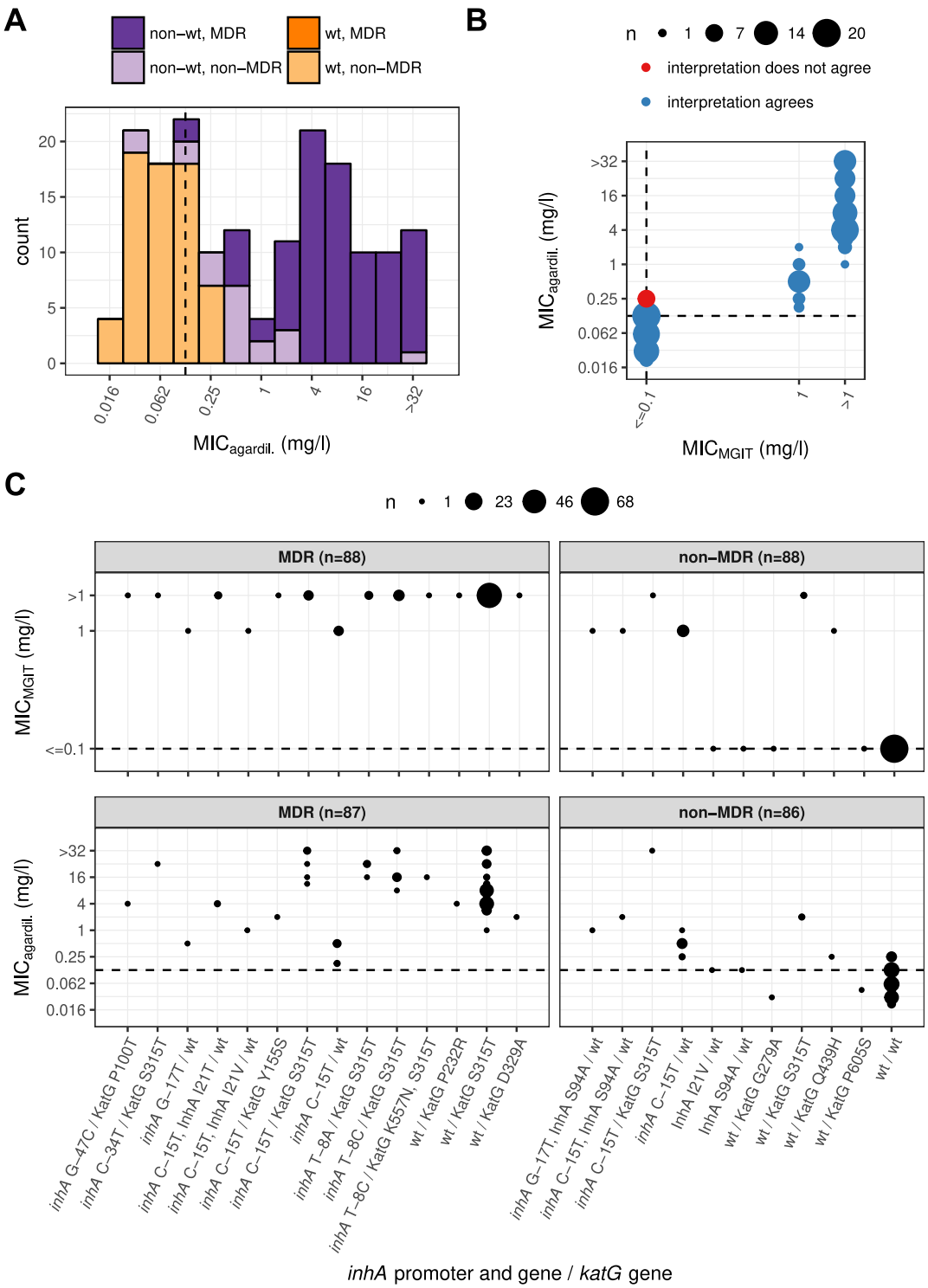

Figure A1.8.: **Summary of DST results for isoniazid.** **A** Histogram of 7H10 agar dilution MICs for isoniazid resistance. **B** method agreement between phenotypic DST by MGIT 960 and 7H10 agar dilution. **C** MICs of *M. tuberculosis* strains harboring resistance mutations in target genes or for wt. Top panel MGIT 960, bottom panel 7H10 agar dilution. In all panels, dashed lines indicate the ECOFF.

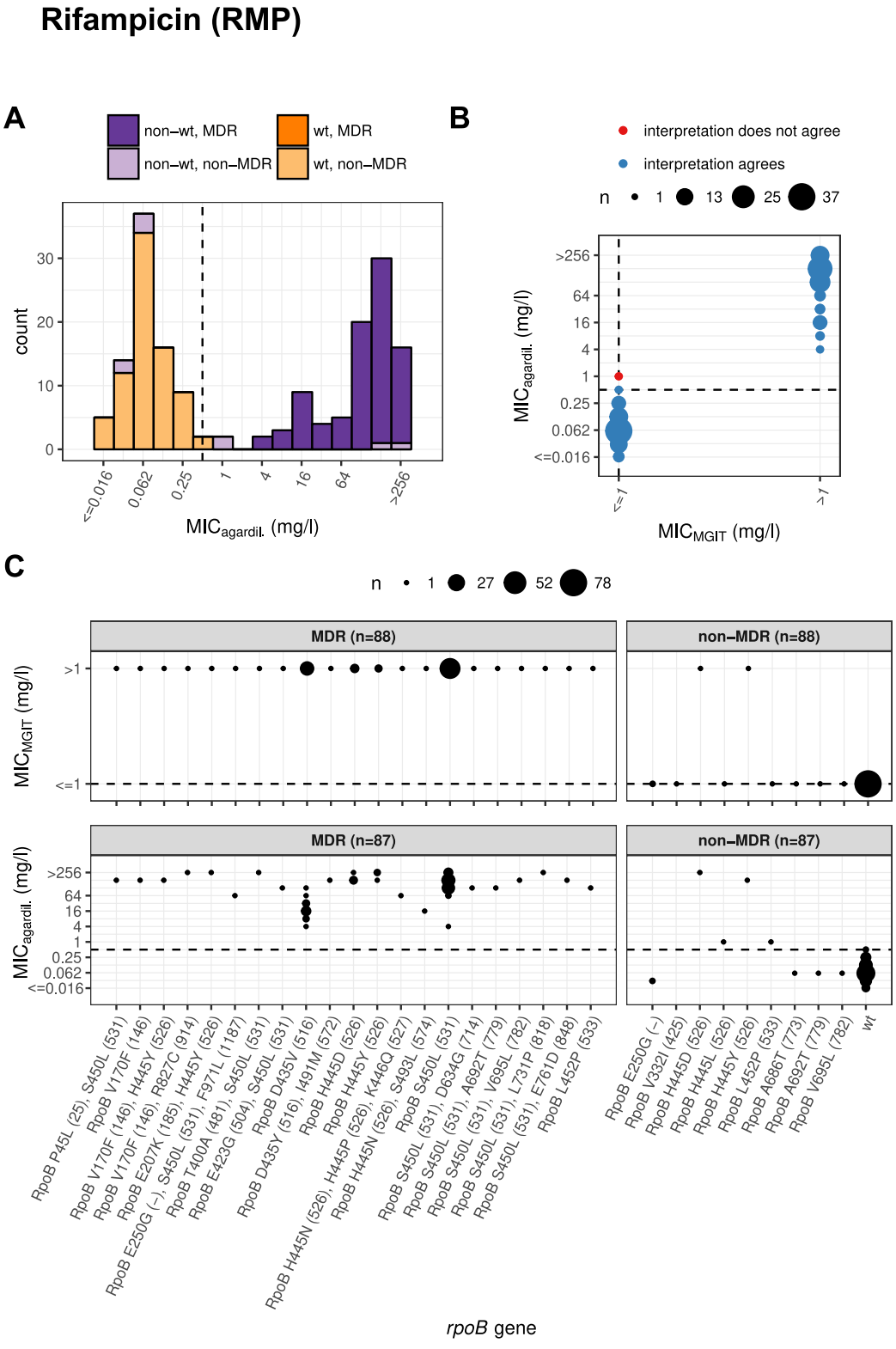

Figure A1.9.: **Summary of DST results for rifampicin.** **A** Histogram of 7H10 agar dilution MICs for rifampicin resistance. **B** method agreement between phenotypic DST by MGIT 960 and 7H10 agar dilution. **C** MICs of *M. tuberculosis* strains harboring resistance mutations in target genes or for wt. Top panel MGIT 960, bottom panel 7H10 agar dilution. Numbers in parentheses indicate the corresponding *Escherichia coli* codon numbers. In all panels, dashed lines indicate the ECOFF (see Table 4.1 in the main text).

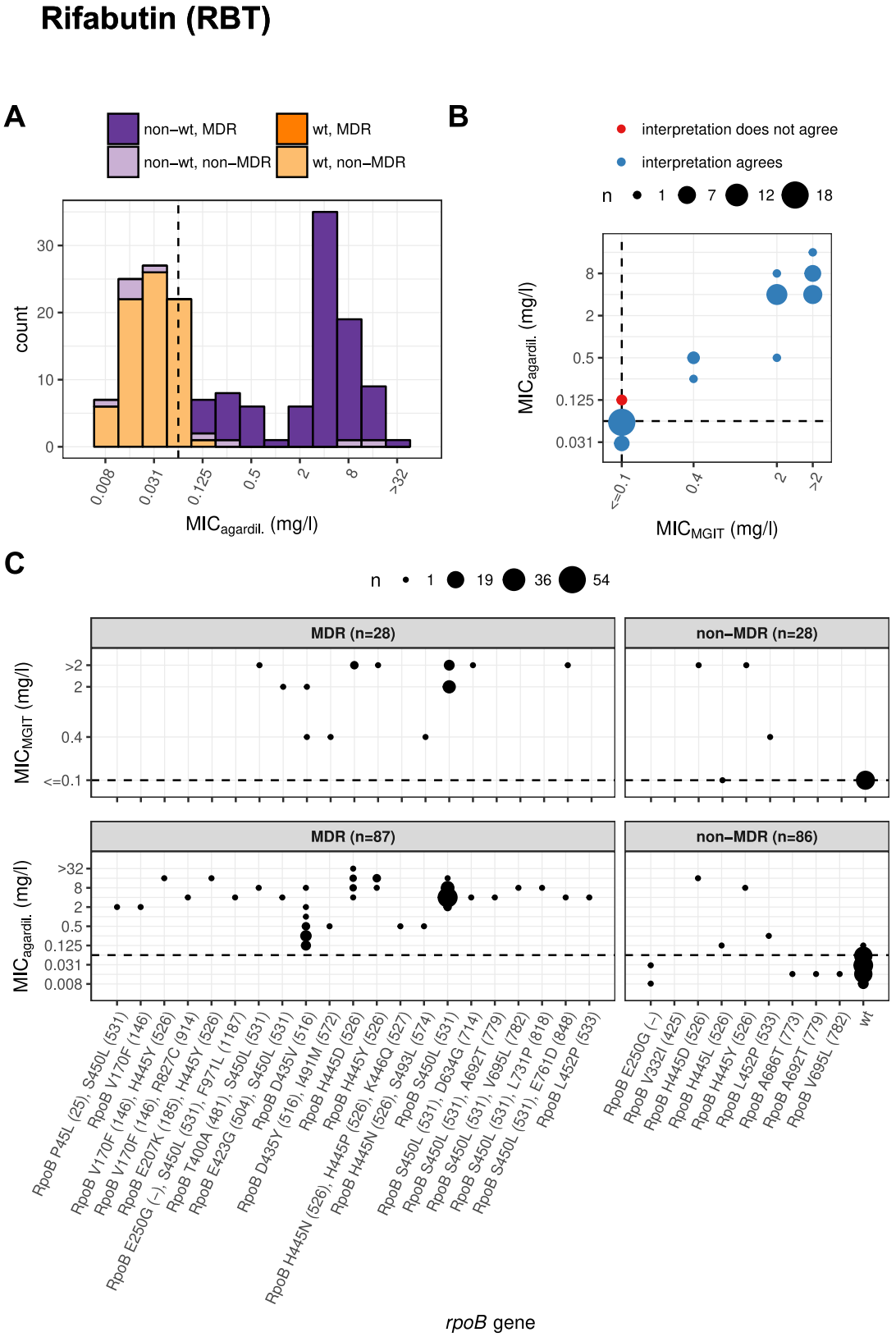

Figure A1.10.: **Summary of DST results for rifabutin. A** Histogram of 7H10 agar dilution MICs for rifabutin resistance. **B** method agreement between phenotypic DST by MGIT 960 and 7H10 agar dilution. **A** MICs of *M. tuberculosis* strains harboring resistance mutations in target genes or for wt. Numbers in parentheses indicate the corresponding *Escherichia coli* codon numbers. Top panel MGIT 960, bottom panel 7H10 agar dilution. In all panels, dashed lines indicate the ECOFF.

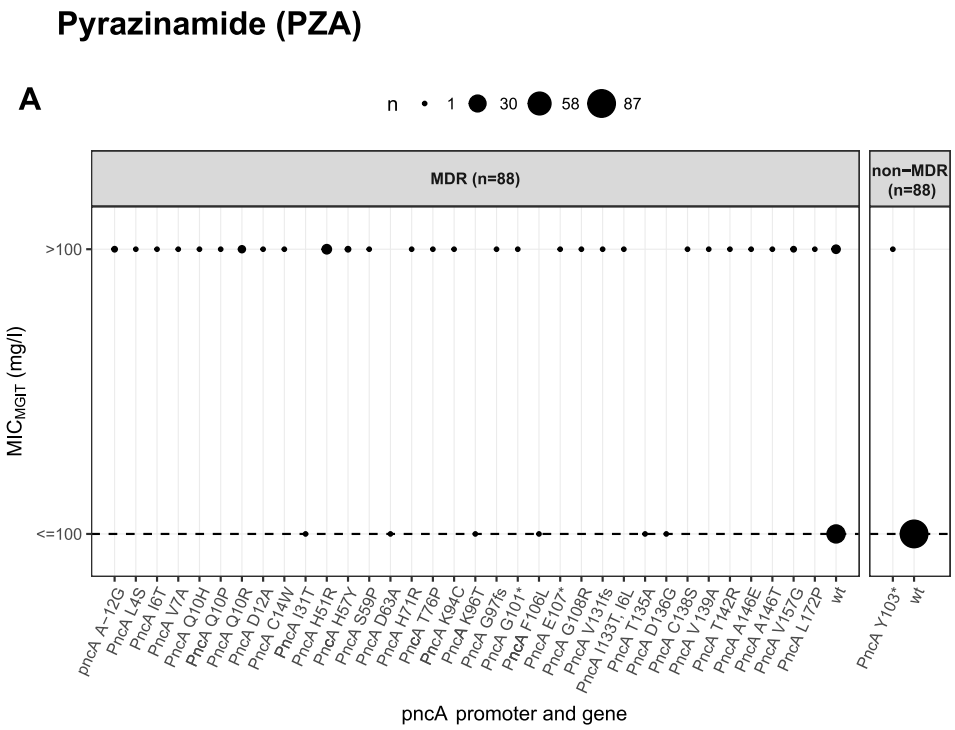

Figure A1.11.: **Summary of DST results for pyrazinamide. A** MICs of *M. tuberculosis* strains harboring resistance mutations in the target gene for pyrazinamide resistance or for wt. Only results for MGIT 960 are available. "*" indicates a stop-gain variant. Dashed line indicates the ECOFF.

A2. Supplement to Chapter 5

A2.1. Materials and methods

Sample set and associated metadata

A retrospective sample (n = 1,003) including all culture-confirmed *Mtb* isolates demonstrating at least an MDR phenotype, collected between 2011 and 2013 and stored by the National Centre for Tuberculosis and Lung Disease (NCTLD) in Tbilisi (Georgia) was re-cultured and processed for DNA extraction for whole genome sequencing. A total of 344 strains were excluded from further analysis due to failed sequencing, no MDR-genotype, large numbers of unfixed positions (potential mixed infections/cross-contamination), or strains with multiple differing metadata entries (Figure A.2.1). The final dataset consisted of 659 whole genome sequences of MDR *Mtb* strains, representing 53 % of all culture-confirmed MDR *Mtb* strains isolated between 2011 and 2013. The per-year sampling coverages were 34.7 % (165/475) in 2012, 82.7 % (286/346) and 52.0 % (208/400) in 2013.The per year lineage proportions remained stable (Figure A.2.2). Limited anonymised patient data including sex, age, incarceration history, current incarceration status, prior TB diagnosis was available. The ethics boards waived the need for individual patient consent on the basis that only limited and anonymised data was provided and that any insight gained from the data would directly be communicated to the public health authorities. For 17.5 % of the analysed strains, no matching metadata was available; for all other strains, complete records were available. The strains without metadata did not group when plotted on the phylogenetic tree (Figure A.2.3.).

Ethical approval

Ethical approval for this study was obtained by the Institutional Review Board of the National Centre for Tuberculosis and Lung Disease in Tbilisi, Georgia and the Ethics Commission of North- and Central Switzerland.

Statistical analyses

Identification of risk factors associated with clustering was performed by multivariable logistic regression. We tested the assumption of linearity between the continuous variables (age and number of drug resistance mutations) and the log odds of the outcome. Fisher's exact and χ^2 tests were used to analyse the association of compensatory mutations with

incarceration and the spill-over of compensated strains to the general public. All analyses were performed using R (v.3.3.3).

Whole genome sequencing

Sequencing libraries were prepared using the Illumina Nextera XT kit and subjected to massive parallel sequencing on the Illumina HiSeq 2500 platform, whereby 125-bp paired-end reads were generated. All sequencing runs were performed at the core sequencing facility of the ETHZ/University of Basel in Basel, Switzerland. The reads were processed using an in-house pipeline as described previously (Ghielmetti *et al.*, 2017) as follows: Trimmomatic (v.0.33) was used to clip adapters and filter for quality, whereby < 20 bp reads were discarded. Overlapping paired-end reads were merged with Seqprep (https://github.com/jstjohn/SeqPrep). The resulting reads were mapped to a reconstructed hypothetical ancestor of the *M. tuberculosis* complex (Comas *et al.*, 2012) with BWA (v.0.7.12), duplicate reads marked with Picard (v.2.1.1) (https://github.com/broadinstitute/picard) with the MarkDuplicates module. To enhance mapping of reads in the vicinity of indels, local realignments with the GATK (v.3.4.0)
modules RealignerTargetCreator and IndelRealigner (McKenna *et al.*, 2010). Pileups were generated with Samtools (v.1.2) (Li *et al.*, 2009b) and single nucleotide variants (SNV) were subsequently called with VarScan (v.2.4.1) (Koboldt *et al.*, 2012) applying the following thresholds: minimum mapping / mimimum base quality of 20, minimum read depth 7x at a given position. SNVs were called if at least 5 reads supported the alternative allele without strand bias. A given SNV was considered fixed if its frequency reached 90 % and a position was called as ancestral if the frequency was below 10 %. The effect of the SNV was inferred using SnpEFF (v.4.11) (Cingolani *et al.*, 2012) using the *M. tuberculosis* H37Rv reference annotation (NC_000962.3). SNVs lying in regions that share ≥ 50 bp sequence identity with other regions in the genome were excluded (Ghielmetti *et al.*, 2017). The unfixed position outliers were defined as having > 3x the interquartile range (IQR) of the ratio between fixed and unfixed positions, whereby the IQR was calculated separately for Lineage 2 and Lineage 4 strains.

Variable position alignment & phylogenetic analysis

Variable SNV pseudo alignments were generated by concatenating all quality filtered SNVs in the dataset. IUPAC nucleotide ambiguity codes were used throughout for unfixed positions, whereby positions were considered variable if at least one isolate in the dataset

had an alternative allele called at a given position. A position was encoded as an X in the alignment, if it was covered by less than 7 reads or if it fell into one of the excluded regions (see above). If a position was not covered, it was encoded as a gap. Two separate alignments were produced: One including genes known to be involved in drug resistance and a separate alignment excluding variable positions in drug resistance-related genes. The former alignment was used for genetic distance-based transmission cluster inference (see below) and the latter was used to infer a maximum likelihood phylogeny using RAxML (v.8.2.8) (Stamatakis, 2014). The phylogeny was inferred using the general time-reversible model of sequence evolution, rooted on *Mycobacterium canettii* (SRR011186). Strains were classified into main- and sub-lineages based on the presence of previously established markers (Coll *et al.*, 2014).

Drug resistance profile prediction

We collated a list of high-confidence drug-resistance mutations (mutations and sources summarised in Table A.2.1.), which we used to screen the sequences. In addition, all non-synonymous substitutions in *ethA*, *pncA*, and Rv0678/*mmpR* were regarded as conferring resistance to ethionamide, pyrazinamide, and bedaquiline, respectively.

Identification of compensatory mutations

We screened the genes *rpoA*, *rpoB* and *rpoC*, encoding the DNA-dependent RNA polymerase, for the presence of non-synonymous mutations. This resulted in a list of rifampicin resistance-conferring, phylogenetic and putative compensatory mutations. Per definition, compensatory mutations must co-occur with rifampicin resistance-conferring mutations but are never found on their own. To remove phylogenetic markers, we collated a list of non-synonymous mutations in *rpoABC* identified in rifampicin susceptible strains (defined as harbouring none of the high-confidence rifampicin resistance-conferring mutations listed in Table A.2.1.) using a large collection of published (Menardo *et al.*, 2018), as well as an in house collection of genomes. After filtering for phylogenetic and rifampicin resistance-conferring mutations, every mutation identified in *rpoABC* was assumed to be a secondary, compensatory mutation. Strains harbouring two mutations affecting the same codon in *rpoB* were assumed to be compensated if one of the mutations conferred rifampicin resistance on its own, e.g. *rpoB* c.1349T>C (Ser450Leu – confers rifampicin resistance) & *rpoB* c.1350G>C (Ser450Ser) combined these two mutations result in the substitution RpoB Ser450Phe.

Genetic distance-based transmission cluster definition

The likelihood of two strains being members of a transmission chain decreases with the number of genetic differences between two strains. Previous analyses have demonstrated that two TB strains isolated from patients with a proven epidemiologic link rarely differ by more than 5 mutations from each other (Walker *et al.*, 2013). A distance matrix based on pairwise SNV distances between any given two strains was inferred using the variable position pseudo alignment including variable positions in drug resistance-related genes with the haplotypes package (v.1.0) for R (v.3.3.3). Insertions/deletions were considered as missing data. Agglomerative clustering was performed using the R package cluster (v.2.0.6) with the agnes function using the unweighted pair-group average method. A threshold of 5 SNV on average was used as a cut-off for likely patient to patient transmission (Walker *et al.*, 2013). The function hclust was used to cut the tree at a height of 5 SNVs. The resulting transmission clusters were filtered for a minimum size of 3 clustered strains, i.e. a minimum of two transmission events to reduce the influence of chance events.

Transmission networks

Transmission graphs were inferred using the R package phybreak (v.0.2.0) (Klinkenberg *et al.*, 2017). Phybreak infers consensus transmission trees by combining transmission models, within-host dynamics, case observation and mutation rate. Phybreak uses Bayesian inference combined with Markov-chain Monte Carlo sampling of the posterior distribution of model parameters, transmission and phylogenetic trees. Priors for the mean of the sampling time ($\mu = 146.28$) and mean of the generation time ($\mu = 173.67$) distributions, as well as the shape parameters (sampling time: $k = 0.021$, generation time: $k = 0.027$) for both distributions were inferred from collated data on the time-course of *M. tuberculosis* infections (Behr *et al.*, 2018) by fitting a gamma distribution with the R package fitdistrplus (v.1.0-11). The mutation rate prior was set at 0.5 mutations per genome and year (Walker *et al.*, 2013). Five independent MCMC chains were run with a burn-in set at 20000 cycles and the sampling of the independent chains was set at 100000 cycles to ensure that most estimated parameters reached an effective sample size > 200 (Lanfear *et al.*, 2016).

A2.2. Supplementary results

Whole genome sequencing summary statistics

The final dataset included 659 strains with a median coverage of 123x (25^{th} percentile = 101, 75^{th} percentile = 141, interquartile range (IQR) = 40). We observed a high percentage of reads that mapped to the reference genome (median = 99.00 %, 25^{th} percentile = 98.73, 75^{th} percentile = 99.09, IQR = 0.36), indicating the absence of major contaminants. The median percentage of the genome with less than 7x coverage was 0.89 % (25^{th} percentile = 0.82, 75^{th} percentile = 0.98, IQR = 0.16). Per genome, a median number of 906 fixed mutations (25^{th} percentile = 900, 75^{th} percentile = 910, IQR = 10) and a median number of 48 unfixed mutations (25^{th} percentile = 43, 75th^{th} percentile = 56, IQR = 13).

A2.3. Supplementary figures

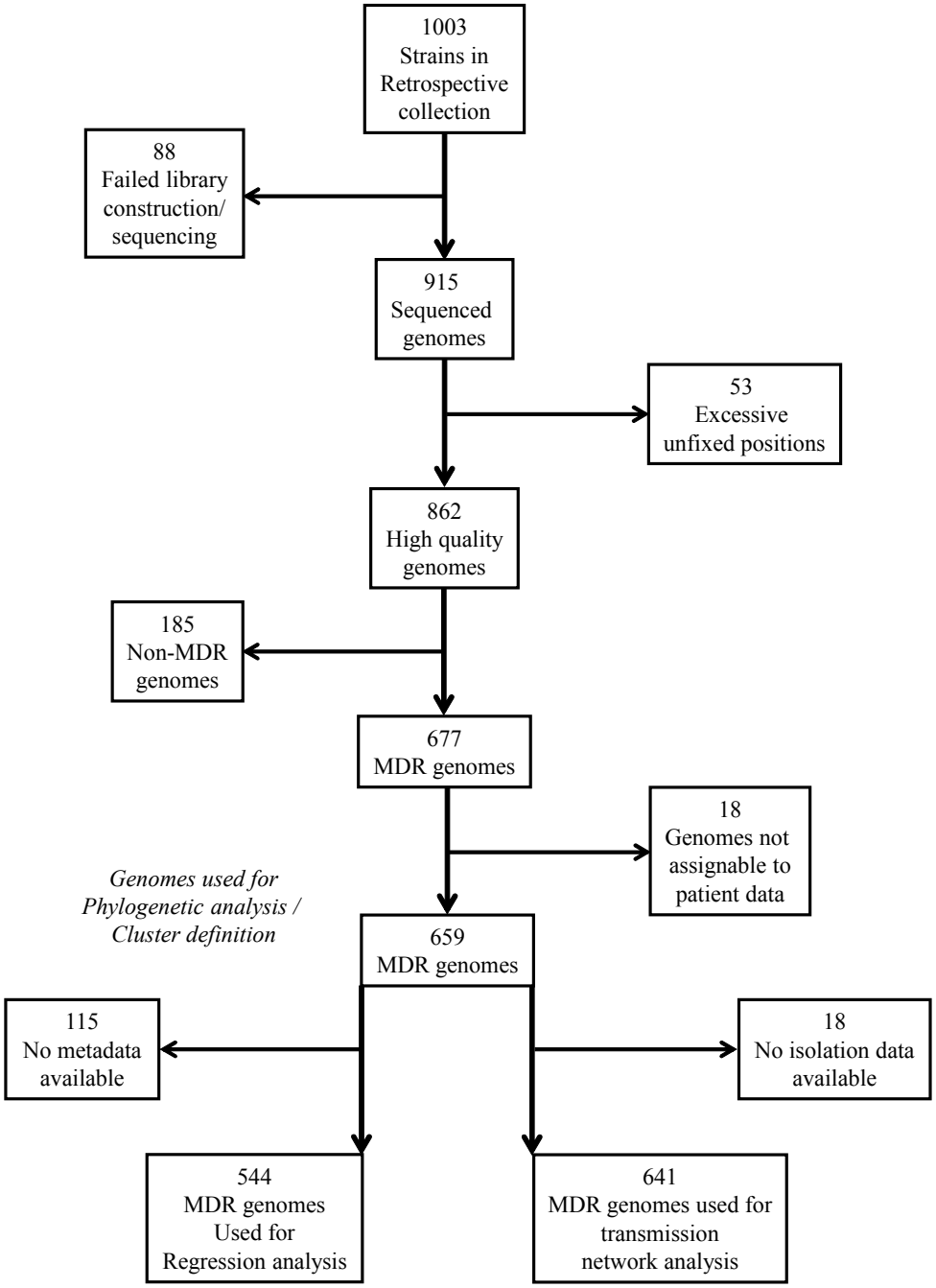

Figure A2.1.: **Flow chart of sample selection.**

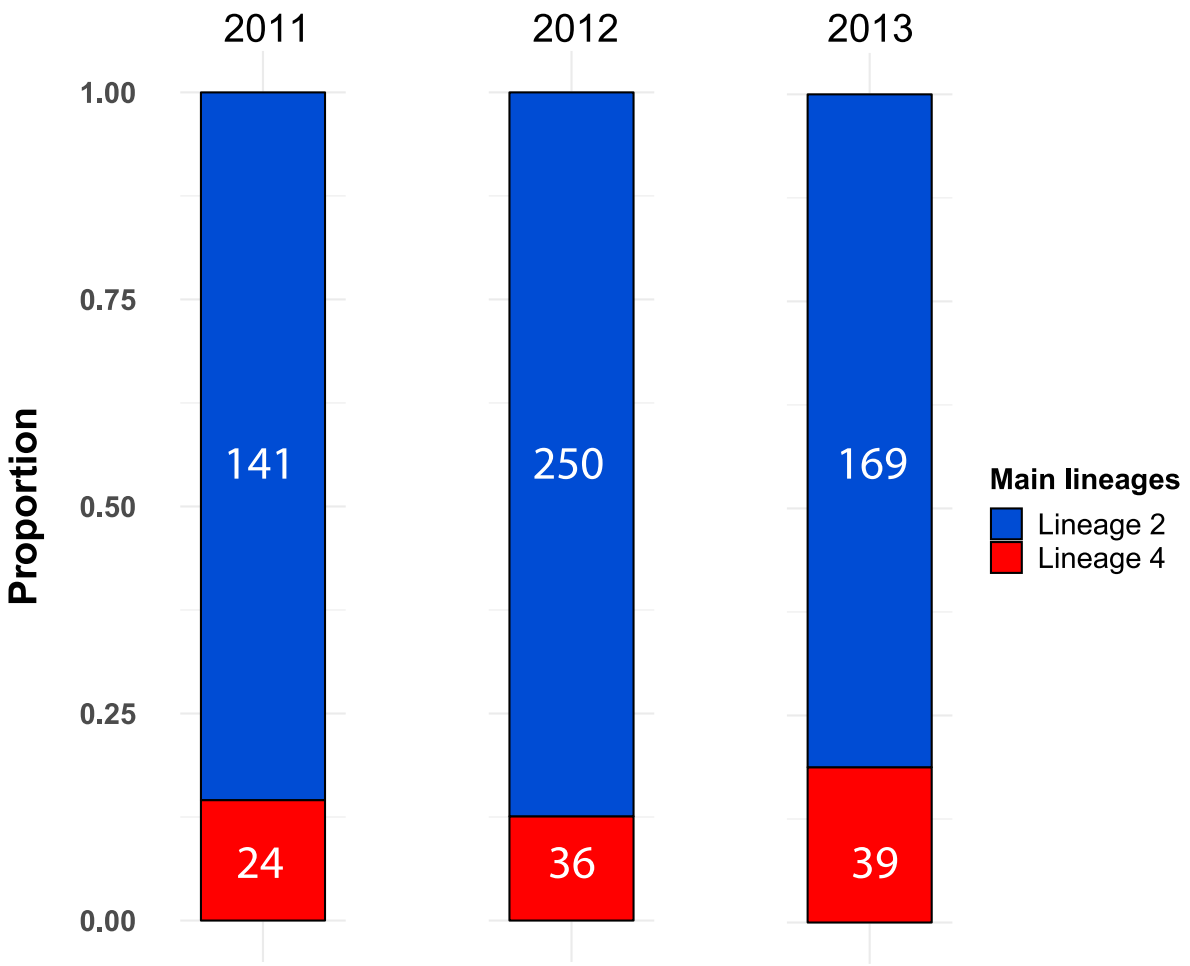

Figure A2.2.: ***M. tuberculosis* Lineage proportions per year.** The proportions of *M. tuberculosis* phylogenetic lineages remained stable during the time frame of the study 2011: 15 % ± 0.05 % Lineage 4, 2012: 13 % ± 0.04 % Lineage 4, 2013:19 % ± 0.06 % Lineage 4; there was no statistical significant difference in proportions per year (χ^2 = 5.27, p = 0.07).

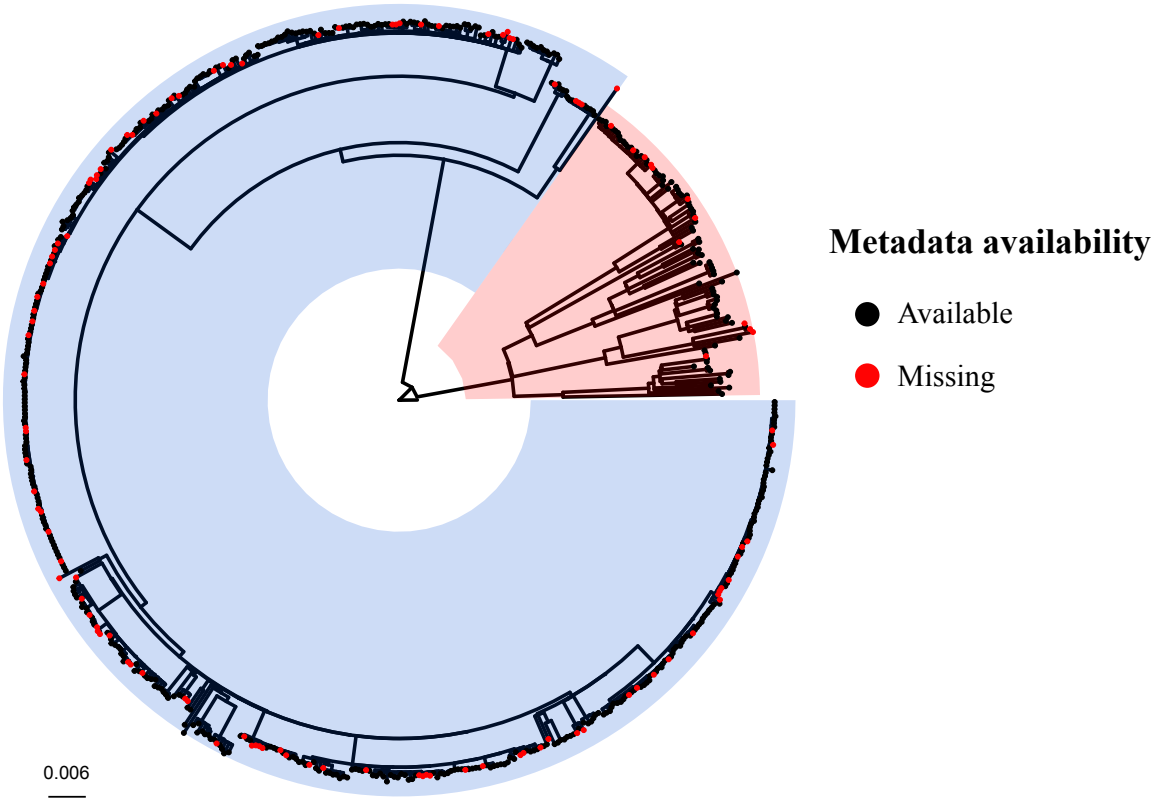

Figure A2.3.: **Maximum likelihood phylogeny of all 659 analysed *M. tuberculosis* genomes.** Lineage 2 strains highlighted in blue, Lineage 4 in red. Scale bar indicates substitutions per site. 115 isolates were discarded as no metadata was available (Figure 1). Availability of metadata for the strains is indicated with coloured tip labels. There is no clustering of strains with missing data. The missing metadata was assumed to be missing completely at random.

A2.4. Supplementary Tables

Table A2.1.: **List of drug resistance mutations compiled from various studies.** Abbreviations of drug names are as follows: SM: Streptomycin, AK: Amikacin, CAP: Capreomycin, KAN: Kanamycin, FQ: Fluoroquinolones, RIF: Rifampicin, INH: Isoniazid, ETH: Ethionamide, PZA: Pyrazinamide, PAS: Para-aminosalicylic acid, EMB: Ethambutol *S: Spontaneous mutant generated in our lab by plating drug sensitive strains on medium containing the corresponding drug.

Genomic position	Ref. base	Alt. base	Locus	Substitution	Drug‡	Ref.
1472358	C	T	rrs	C513T	SM	(Walker *et al.*, 2015)
1472359	A	C	rrs	A514C	SM	(Walker *et al.*, 2015)
1472361	C	T	rrs	C516T	SM	(Tudo *et al.*, 2010)
1472362	C	T	rrs	C517T	SM	(Walker *et al.*, 2015)
1472749	T	G	rrs	C904G	SM	(Honore *et al.*, 1994)
1472749	T	A	rrs	C904A	SM	(Honore *et al.*, 1994)
1472751	A	G	rrs	A906G	SM	(Jagielski *et al.*, 2014)
1472752	A	T	rrs	A907T	SM	(Jagielski *et al.*, 2014)
1473246	A	G	rrs	A1401G	AK;CAP;KAN	(Walker *et al.*, 2015)
1473247	C	T	rrs	C1402T	CAP	(Maus *et al.*, 2005)
1473329	G	T	rrs	G1484T	AK;CAP	(Walker *et al.*, 2015)
1472359	A	T	rrs	A514T	SM	(Walker *et al.*, 2015)
1472750	C	G	rrs	C905G	SM	(Gygli *et al.*, 2018)
6620	G	A	Rv0005	D461N	FQ	(Malik *et al.*, 2012)
6734	A	G	Rv0005	N499D	FQ	(Malik *et al.*, 2012)
7563	G	T	Rv0006	G88C	FQ	S*
7564	G	C	Rv0006	G88A	FQ	(Matrat *et al.*, 2006)
7570	C	T	Rv0006	A90V	FQ	(Walker *et al.*, 2015)

Table A2.1 continued from previous page

Genomic position	Ref. base	Alt. base	Locus	Substitution	Drug‡	Ref.
7572	T	C	Rv0006	S91P	FQ	(Walker *et al.*, 2015)
7581	G	A	Rv0006	D94N	FQ	(Walker *et al.*, 2015)
7581	G	C	Rv0006	D94H	FQ	S*
7581	G	T	Rv0006	D94Y	FQ	(Bloemberg *et al.*, 2015)
7582	A	G	Rv0006	D94G	FQ	(Walker *et al.*, 2015)
7582	A	C	Rv0006	D94A	FQ	(Walker *et al.*, 2015)
760314	G	T	Rv0667	V170F	RIF	(Walker *et al.*, 2015)
760882	T	C	Rv0667	V359A	RIF	(Walker *et al.*, 2015)
761100	C	G	Rv0667	Q432E	RIF	S*
761100	C	A	Rv0667	Q432K	RIF	(Walker *et al.*, 2015)
761101	A	C	Rv0667	Q432P	RIF	(Miotto *et al.*, 2017)
761101	A	T	Rv0667	Q432L	RIF	S*
761109	G	T	Rv0667	D435Y	RIF	S*
761110	A	G	Rv0667	D435G	RIF	(Miotto *et al.*, 2017)
761110	A	T	Rv0667	D435V	RIF	(Walker *et al.*, 2015)
761128	C	T	Rv0667	S441L	RIF	S*
761139	C	G	Rv0667	H445D	RIF	(Walker *et al.*, 2015)
761139	C	T	Rv0667	H445Y	RIF	(Walker *et al.*, 2015)
761139	C	A	Rv0667	H445N	RIF	(Walker *et al.*, 2015)
761140	A	T	Rv0667	H445L	RIF	(Miotto *et al.*, 2017)
761140	A	C	Rv0667	H445P	RIF	(Gagneux *et al.*, 2006)
761140	A	G	Rv0667	H445R	RIF	(Walker *et al.*, 2015)

Table A2.1 continued from previous page

Genomic position	Ref. base	Alt. base	Locus	Substitution	Drug‡	Ref.
761149	G	A	Rv0667	R448Q	RIF	S*
761155	C	T	Rv0667	S450L	RIF	(Walker *et al.*, 2015)
761155	C	G	Rv0667	S450W	RIF	(Walker *et al.*, 2015)
761161	T	C	Rv0667	L452P	RIF	(Walker *et al.*, 2015)
761277	A	T	Rv0667	I491F	RIF	(Walker *et al.*, 2015)
761155/761156	C/G	T/C	Rv0667	S450F	RIF	(Walker *et al.*, 2015)
761139/761140	C/A	T/T	Rv0667	H445S	RIF	(Miotto *et al.*, 2017)
761139/761140	C/A	T/G	Rv0667	H445C	RIF	(Miotto *et al.*, 2017)
761109/761110	G/A	T/T	Rv0667	D435F	RIF	(Walker *et al.*, 2015)
761095	T	C	Rv0667	L430P	RIF	(Miotto *et al.*, 2017)
761095	T	G	Rv0667	L430R	RIF	(Miotto *et al.*, 2017)
761277	A	G	Rv0667	I491V	RIF	(Walker *et al.*, 2015)
781687	A	G	Rv0682	K43R	SM	(Walker *et al.*, 2015)
781822	A	G	Rv0682	K88R	SM	(Walker *et al.*, 2015)
1673423	G	T	Rv1483	G-17T	INH;ETH	(Walker *et al.*, 2015)
1673425	C	T	Rv1483	C-15T	INH;ETH	(Walker *et al.*, 2015)
1673432	T	C	Rv1483	T-8C	INH;ETH	(Walker *et al.*, 2015)
1674263	T	C	Rv1484	I21T	INH;ETH	(Walker *et al.*, 2015)
1674481	T	G	Rv1484	S94A	INH;ETH	(Walker *et al.*, 2015)
1674782	T	C	Rv1484	I194T	INH	(Walker *et al.*, 2015)
1917857	C	T	Rv1694	C-83T	CAP	(Walker *et al.*, 2015)
1918647	T	A	Rv1694	N236K	CAP	(Miotto *et al.*, 2017)

Table A2.1 continued from previous page

Genomic position	Ref. base	Alt. base	Locus	Substitution	Drug‡	Ref.
1918647	T	G	Rv1694	N236K	CAP	(Miotto *et al.*, 2017)
2154214	A	G	Rv1908c	V633A	INH	(Walker *et al.*, 2015)
2154218	G	A	Rv1908c	R632C	INH	(Ando *et al.*, 2010)
2154488	C	G	Rv1908c	D542H	INH	(Ando *et al.*, 2010)
2154853	A	G	Rv1908c	M420T	INH	(Ando *et al.*, 2010)
2154857	C	G	Rv1908c	D419H	INH	(Ando *et al.*, 2010)
2155059	C	T	Rv1908c	W351STOP	INH	S*
2155060	C	T	Rv1908c	W351STOP	INH	S*
2155126	T	G	Rv1908c	D329A	INH	(Gygli *et al.*, 2018)
2155129	C	A	Rv1908c	W328L	INH	(Walker *et al.*, 2015)
2155167	G	T	Rv1908c	S315R	INH	(Ando *et al.*, 2010)
2155167	G	C	Rv1908c	S315R	INH	(Ando *et al.*, 2010)
2155168	C	A	Rv1908c	S315I	INH	S*
2155168	C	G	Rv1908c	S315T	INH	(Walker *et al.*, 2015)
2155168	C	T	Rv1908c	S315N	INH	(Walker *et al.*, 2015)
2155169	T	G	Rv1908c	S315R	INH	(Ando *et al.*, 2010)
2155212	C	A	Rv1908c	W300C	INH	(Walker *et al.*, 2015)
2155212	C	G	Rv1908c	W300C	INH	(Walker *et al.*, 2015)
2155289	T	G	Rv1908c	T275P	INH	(Pym *et al.*, 2002)
2155417	G	C	Rv1908c	P232R	INH	(Gygli *et al.*, 2018)
2155541	A	T	Rv1908c	W191R	INH	(Walker *et al.*, 2015)
2155541	A	G	Rv1908c	W191R	INH	(Walker *et al.*, 2015)

Table A2.1 continued from previous page

Genomic position	Ref. base	Alt. base	Locus	Substitution	Drug‡	Ref.
2155573	G	T	Rv1908c	T180K	INH	(Walker *et al.*, 2015)
2155636	A	G	Rv1908c	L159P	INH	(Walker *et al.*, 2015)
2155648	T	G	Rv1908c	Y155S	INH	(Gygli *et al.*, 2018)
2155814	G	T	Rv1908c	P100T	INH	(Gygli *et al.*, 2018)
2155844	A	T	Rv1908c	W90R	INH	(Walker *et al.*, 2015)
2155844	A	G	Rv1908c	W90R	INH	(Walker *et al.*, 2015)
2155169	T	C	Rv1908c	S315G	INH	(Suarez *et al.*, 2009)
2289252	T	C	Rv2043c	A-11G	PZA	(Walker *et al.*, 2015)
2289253	A	G	Rv2043c	T-12C	PZA	(Walker *et al.*, 2015)
2289252	T	G	Rv2043c	A-11C	PZA	(Ramirez-Busby *et al.*, 2015)
2715338	C	A	Rv2416c	G-6T	KAN	(Engström *et al.*, 2011)
2715342	C	T	Rv2416c	G-10A	KAN	(Walker *et al.*, 2015)
2715342	C	G	Rv2416c	G-10C	KAN	(Kambli *et al.*, 2016)
2715344	G	A	Rv2416c	C-12T	KAN	(Kambli *et al.*, 2016)
2715345	T	C	Rv2416c	A-13G	KAN	(Zaunbrecher *et al.*, 2009)
2715346	G	A	Rv2416c	C-14T	KAN;CAP	(Miotto *et al.*, 2017)
2715347	G	A	Rv2416c	C-15T	KAN	(Engström *et al.*, 2011)
2715369	C	A	Rv2416c	G-37T	KAN	(Zaunbrecher *et al.*, 2009)
2726136	C	T	Rv2428	C-57T	INH	(Walker *et al.*, 2015)
2747141	T	C	Rv2447c	E153G	PAS	(Coll *et al.*, 2015)
2747144	A	G	Rv2447c	F152S	PAS	(Coll *et al.*, 2015)
2747145	A	G	Rv2447c	F152L	PAS	(Coll *et al.*, 2015)

Table A2.1 continued from previous page

Genomic position	Ref. base	Alt. base	Locus	Substitution	Drug‡	Ref.
2747149	G	C	Rv2447c	S150R	PAS	(Coll *et al.*, 2015)
2747151	T	C	Rv2447c	S150G	PAS	(Coll *et al.*, 2015)
2747381	T	C	Rv2447c	N73S	PAS	(Coll *et al.*, 2015)
2747480	T	C	Rv2447c	E40G	PAS	(Coll *et al.*, 2015)
2747052	C	G	Rv2447c	A183P	PAS	(Coll *et al.*, 2015)
2747454	G	A	Rv2447c	R49W	PAS	(Coll *et al.*, 2015)
2747541	T	G	Rv2447c	T20P	PAS	(Coll *et al.*, 2015)
2747481	C	G	Rv2447c	E40Q	PAS	(Coll *et al.*, 2015)
2747471	A	G	Rv2447c	I43T	PAS	(Coll *et al.*, 2015)
2747471	A	C	Rv2447c	I43S	PAS	(Coll *et al.*, 2015)
2747453	C	G	Rv2447c	R49P	PAS	(Coll *et al.*, 2015)
2747433	G	C	Rv2447c	L56V	PAS	(Coll *et al.*, 2015)
2747328	G	A	Rv2447c	R91W	PAS	(Coll *et al.*, 2015)
2747151	T	A	Rv2447c	S150C	PAS	(Coll *et al.*, 2015)
2747141	T	G	Rv2447c	E153A	PAS	(Coll *et al.*, 2015)
2986860	G	C	Rv2671	G8R	PAS	(Coll *et al.*, 2015)
3073808	G	C	Rv2764c	R222G	PAS	(Coll *et al.*, 2015)
3074171	A	G	Rv2764c	W101R	PAS	(Coll *et al.*, 2015)
3074249	G	T	Rv2764c	H75N	PAS	(Coll *et al.*, 2015)
3074408	T	C	Rv2764c	T22A	PAS	(Coll *et al.*, 2015)
3074365	T	C	Rv2764c	Y36C	PAS	(Coll *et al.*, 2015)
3074246	C	A	Rv2764c	G76*	PAS	(Coll *et al.*, 2015)

Table A2.1 continued from previous page

Genomic position	Ref. base	Alt. base	Locus	Substitution	Drug‡	Ref.
3074243	C	A	Rv2764c	V77F	PAS	(Coll *et al.*, 2015)
3074223	C	A	Rv2764c	W83C	PAS	(Coll *et al.*, 2015)
3074224	C	T	Rv2764c	W83*	PAS	(Coll *et al.*, 2015)
3074201	C	T	Rv2764c	G91R	PAS	(Coll *et al.*, 2015)
3074178	C	T	Rv2764c	W98*	PAS	(Coll *et al.*, 2015)
3074159	A	G	Rv2764c	S105P	PAS	(Coll *et al.*, 2015)
3074095	C	T	Rv2764c	R126Q	PAS	(Coll *et al.*, 2015)
3074018	A	C	Rv2764c	F152V	PAS	(Coll *et al.*, 2015)
3073990	C	T	Rv2764c	C161Y	PAS	(Coll *et al.*, 2015)
3073852	T	C	Rv2764c	H207R	PAS	(Coll *et al.*, 2015)
3073801	G	A	Rv2764c	P224L	PAS	(Coll *et al.*, 2015)
3073768	C	G	Rv2764c	R235P	PAS	(Coll *et al.*, 2015)
3074211	TG	T	Rv2764c	NA	PAS	(Coll *et al.*, 2015)
3074360	AA	A	Rv2764c	NA	PAS	(Coll *et al.*, 2015)
3074254	T	TGTGCTCGTG	Rv2764c	NA	PAS	(Coll *et al.*, 2015)
3074099	A	A	Rv2764c	NA	PAS	(Coll *et al.*, 2015)
3073999	CG	C	Rv2764c	NA	PAS	(Coll *et al.*, 2015)
4243217	C	G	Rv3794	C-16G	EMB	(Walker *et al.*, 2015)
4243221	C	T	Rv3794	C-12T	EMB	(Walker *et al.*, 2015)
4243217	C	T	Rv3794	C-16T	EMB	(Cui *et al.*, 2014)
4243217	C	A	Rv3794	C-16A	EMB	(Cui *et al.*, 2014)
4247429	A	T	Rv3795	M306L	EMB	(Safi *et al.*, 2013)

Table A2.1 continued from previous page

Genomic position	Ref. base	Alt. base	Locus	Substitution	Drug‡	Ref.
4247429	A	C	Rv3795	M306L	EMB	(Safi *et al.*, 2013)
4247429	A	G	Rv3795	M306V	EMB	(Walker *et al.*, 2015)
4247431	G	C	Rv3795	M306I	EMB	(Walker *et al.*, 2015)
4247431	G	A	Rv3795	M306I	EMB	(Walker *et al.*, 2015)
4247431	G	T	Rv3795	M306I	EMB	(Walker *et al.*, 2015)
4247469	A	C	Rv3795	Y319S	EMB	(Gygli *et al.*, 2018)
4247495	G	T	Rv3795	D328Y	EMB	(Gygli *et al.*, 2018)
4247496	A	G	Rv3795	D328G	EMB	(Gygli *et al.*, 2018)
4247574	A	C	Rv3795	D354A	EMB	(Walker *et al.*, 2015)
4247728	G	T	Rv3795	E405D	EMB	(Gygli *et al.*, 2018)
4247728	G	C	Rv3795	E405D	EMB	(Gygli *et al.*, 2018)
4247729	G	T	Rv3795	G406C	EMB	(Safi *et al.*, 2013)
4247730	G	A	Rv3795	G406D	EMB	(Gygli *et al.*, 2018)
4247730	G	C	Rv3795	G406A	EMB	(Gygli *et al.*, 2018)
4248002	C	A	Rv3795	Q497K	EMB	(Walker *et al.*, 2015)
4248003	A	T	Rv3795	Q497L	EMB	(Gygli *et al.*, 2018)
4248003	A	G	Rv3795	Q497R	EMB	(Walker *et al.*, 2015)
4248027	C	T	Rv3795	A505V	EMB	(Gygli *et al.*, 2018)
4249512	T	G	Rv3795	M1000R	EMB	(Gygli *et al.*, 2018)
4249583	G	A	Rv3795	D1024N	EMB	(Gygli *et al.*, 2018)
4249757	A	G	Rv3795	T1082A	EMB	(Gygli *et al.*, 2018)
4248003	A	C	Rv3795	Q497P	EMB	(Gygli *et al.*, 2018)

Table A2.1 continued from previous page

Genomic position	Ref. base	Alt. base	Locus	Substitution	Drug‡	Ref.
4247469	A	G	Rv3795	Y319C	EMB	(Gygli *et al.*, 2018)
4407554	G	C	Rv3919c	R217G	SM	(Gygli *et al.*, 2018)
4407598	A	T	Rv3919c	A202E	SM	(Walker *et al.*, 2015)
4407604	G	T	Rv3919c	A200E	SM	(Walker *et al.*, 2015)
4407730	C	A	Rv3919c	R158L	SM	(Gygli *et al.*, 2018)
4407756	G	T	Rv3919c	S149R	SM	(Gygli *et al.*, 2018)
4407756	G	C	Rv3919c	S149R	SM	(Gygli *et al.*, 2018)
4407758	T	G	Rv3919c	S149R	SM	(Gygli *et al.*, 2018)
4407790	G	A	Rv3919c	A138V	SM	(Walker *et al.*, 2015)
4407791	C	T	Rv3919c	A138T	SM	(Walker *et al.*, 2015)
4407794	G	A	Rv3919c	R137W	SM	(Walker *et al.*, 2015)
4407802	G	T	Rv3919c	A134E	SM	(Walker *et al.*, 2015)
4407880	A	C	Rv3919c	L108R	SM	(Gygli *et al.*, 2018)
4407922	A	G	Rv3919c	L94P	SM	(Walker *et al.*, 2015)
4407931	A	G	Rv3919c	L91P	SM	(Walker *et al.*, 2015)
4407940	A	G	Rv3919c	V88A	SM	(Walker *et al.*, 2015)
4407952	G	A	Rv3919c	P84L	SM	(Gygli *et al.*, 2018)
4407965	C	G	Rv3919c	A80P	SM	(Walker *et al.*, 2015)
4407973	A	C	Rv3919c	V77G	SM	(Gygli *et al.*, 2018)
4407979	G	A	Rv3919c	P75L	SM	(Walker *et al.*, 2015)
4407985	C	T	Rv3919c	G73E	SM	(Gygli *et al.*, 2018)
4407985	C	G	Rv3919c	G73A	SM	(Gygli *et al.*, 2018)

Table A2.1 continued from previous page

Genomic position	Ref. base	Alt. base	Locus	Substitution	Drug‡	Ref.
4407994	C	T	Rv3919c	S70N	SM	(Walker *et al.*, 2015)
4407997	C	T	Rv3919c	G69D	SM	(Walker *et al.*, 2015)
4408003	T	C	Rv3919c	D67G	SM	(Gygli *et al.*, 2018)
4408007	C	A	Rv3919c	V66L	SM	(Gygli *et al.*, 2018)
4408007	C	G	Rv3919c	V66L	SM	(Gygli *et al.*, 2018)
4408009	A	C	Rv3919c	V65G	SM	(Walker *et al.*, 2015)
4408061	G	T	Rv3919c	H48N	SM	(Walker *et al.*, 2015)
4407952	G	C	Rv3919c	P84R	SM	(Gygli *et al.*, 2018)

Supplementary References

Ando, H, Kondo, Y, Suetake, T, Toyota, E, Kato, S, Mori, T, Kirikae, T (2010) Identification of katG Mutations Associated with High-Level Isoniazid Resistance in Mycobacterium tuberculosis. *Antimicrobial Agents and Chemotherapy*, **54**(5): 1793–1799.

Ängeby, K, Juréen, P, Kahlmeter, G, Hoffner, S, Schön, T (2012) Challenging a dogma: antimicrobial susceptibility testing breakpoints for Mycobacterium tuberculosis. *Bulletin of the World Health Organization*, **90**(9): 693–698.

Behr, MA, Edelstein, PH, Ramakrishnan, L (2018) Revisiting the timetable of tuberculosis. *BMJ*, **2738**(August): k2738.

Bloemberg, GV, Keller, PM, Stucki, D, Trauner, A, Borrell, S, Latshang, T, Coscolla, M, Rothe, T, Hömke, R, Ritter, C, Feldmann, J, Schulthess, B, Gagneux, S, Böttger, EC (2015) Acquired Resistance to Bedaquiline and Delamanid in Therapy for Tuberculosis. *New England Journal of Medicine*, **373**(20): 1986–1988.

Cingolani, P, Platts, A, Wang, LL, Coon, M, Nguyen, T, Wang, L, Land, SJ, Lu, X, Ruden, DM (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly*, **6**(2): 80–92.

Coll, F, McNerney, R, Guerra-Assunção, JA, Glynn, JR, Perdigão, J, Viveiros, M, Portugal, I, Pain, A, Martin, N, Clark, TG (2014) A robust SNP barcode for typing Mycobacterium tuberculosis complex strains. *Nature Communications*, **5**(1): 4812.

Coll, F, McNerney, R, Preston, MD, Guerra-Assunção, JA, Warry, A, Hill-Cawthorne, G, Mallard, K, Nair, M, Miranda, A, Alves, A, Perdigão, J, Viveiros, M, Portugal, I, Hasan, Z, Hasan, R, Glynn, JR, Martin, N, Pain, A, Clark, TG (2015) Rapid determination of anti-tuberculosis drug resistance from whole-genome sequences. *Genome Medicine*, **7**(1): 51.

Comas, I, Borrell, S, Roetzer, A, Rose, G, Malla, B, Kato-Maeda, M, Galagan, J, Niemann, S, Gagneux, S (2012) Whole-genome sequencing of rifampicin-resistant Mycobacterium tuberculosis strains identifies compensatory mutations in RNA polymerase genes. *Nature Genetics*, **44**(1): 106–110.

Comas, I, Chakravartti, J, Small, PM, Galagan, J, Niemann, S, Kremer, K, Ernst, JD, Gagneux, S (2010) Human T cell epitopes of Mycobacterium tuberculosis are evolutionarily hyperconserved. *Nature Genetics*, **42**(6): 498–503.

Cui, Z, Li, Y, Cheng, S, Yang, H, Lu, J, Hu, Z, Ge, B (2014) Mutations in the embC-embA intergenic region contribute to mycobacterium tuberculosis resistance to ethambutol. *Antimicrobial Agents and Chemotherapy*, **58**(11): 6837–6843.

Engström, A, Perskvist, N, Werngren, J, Hoffner, SE, Juréen, P (2011) Comparison of clinical isolates and in vitro selected mutants reveals that tlyA is not a sensitive genetic

marker for capreomycin resistance in Mycobacterium tuberculosis. *Journal of Antimicrobial Chemotherapy*, **66**(6): 1247–1254.

Gagneux, S, DeRiemer, K, Van, T, Kato-Maeda, M, Jong, BC de, Narayanan, S, Nicol, M, Niemann, S, Kremer, K, Gutierrez, MC, Hilty, M, Hopewell, PC, Small, PM (2006) Variable host-pathogen compatibility in Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences of the United States of America*, **103**(8): 2869–73.

Ghielmetti, G, Coscolla, M, Ruetten, M, Friedel, U, Loiseau, C, Feldmann, J, Steinmetz, HW, Stucki, D, Gagneux, S (2017) Tuberculosis in Swiss captive Asian elephants: microevolution of Mycobacterium tuberculosis characterized by multilocus variable-number tandem-repeat analysis and whole-genome sequencing. *Scientific Reports*, **7**(1): 14647.

Gygli, SM, Borrell, S, Trauner, A, Gagneux, S (2017) Antimicrobial resistance in Mycobacterium tuberculosis: mechanistic and evolutionary perspectives. *FEMS Microbiology Reviews*, **41**(3): 354–373.

Gygli, SM, Keller, PM, Ballif, M, Blöchliger, N, Hömke, R, Reinhard, M, Loiseau, C, Ritter, C, Sander, P, Borrell, S, Loo, JC, Avihingsanon, A, Gnokoro, J, Yotebieng, M, Egger, M, Gagneux, S, Böttger, EC (2018) Whole genome sequencing for drug resistance profile prediction in Mycobacterium tuberculosis. *bioRxiv*, (401703):

Honore, N, Cole, ST (1994) Streptomycin resistance in mycobacteria. *Antimicrobial Agents and Chemotherapy*, **38**(2): 238–242.

Jagielski, T, Ignatowska, H, Bakuła, Z, Dziewit, Ł, Napiórkowska, A, Augustynowicz-Kopeć, E, Zwolska, Z, Bielecki, J (2014) Screening for Streptomycin Resistance-Conferring Mutations in Mycobacterium tuberculosis Clinical Isolates from Poland. *PLoS ONE*, **9**(6): ed. by I Mokrousov, e100078.

Kambli, P, Ajbani, K, Nikam, C, Sadani, M, Shetty, A, Udwadia, Z, Georghiou, SB, Rodwell, TC, Catanzaro, A, Rodrigues, C (2016) Correlating rrs and eis promoter mutations in clinical isolates of Mycobacterium tuberculosis with phenotypic susceptibility levels to the second-line injectables. *International Journal of Mycobacteriology*, **5**(1): 1–6.

Klinkenberg, D, Backer, JA, Didelot, X, Colijn, C, Wallinga, J (2017) Simultaneous inference of phylogenetic and transmission trees in infectious disease outbreaks. *PLOS Computational Biology*, **13**(5): ed. by MM Tanaka, e1005495.

Koboldt, DC, Zhang, Q, Larson, DE, Shen, D, McLellan, MD, Lin, L, Miller, Ca, Mardis, ER, Ding, L, Wilson, RK (2012) VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Research*, **22**(3): 568–576.

Lanfear, R, Hua, X, Warren, DL (2016) Estimating the effective sample size of tree topologies from Bayesian phylogenetic analyses. *Genome Biology and Evolution*, **8**(8): 2319–2332.

Li, H, Durbin, R (2009a) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**(14): 1754–1760. eprint: 1303.3997.

Li, H, Handsaker, B, Wysoker, A, Fennell, T, Ruan, J, Homer, N, Marth, G, Abecasis, G, Durbin, R (2009b) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**(16): 2078–2079. eprint: 1006.1266v2.

Malik, S, Willby, M, Sikes, D, Tsodikov, OV, Posey, JE (2012) New Insights into Fluoroquinolone Resistance in Mycobacterium tuberculosis: Functional Genetic Analysis of gyrA and gyrB Mutations. *PLoS ONE*, **7**(6): e39754.

Matrat, S, Veziris, N, Mayer, C, Jarlier, V, Truffot-Pernot, C, Camuset, J, Bouvet, E, Cambau, E, Aubry, A (2006) Functional analysis of DNA gyrase mutant enzymes carrying mutations at position 88 in the A subunit found in clinical strains of Mycobacterium tuberculosis resistant to fluoroquinolones. *Antimicrobial Agents and Chemotherapy*, **50**(12): 4170–4173.

Maus, CE, Plikaytis, BB, Shinnick, TM (2005) Mutation of tlyA Confers Capreomycin Resistance in Mycobacterium tuberculosis. *Antimicrobial Agents and Chemotherapy*, **49**(2): 571–577.

McKenna, A, Hanna, M, Banks, E, Sivachenko, A, Cibulskis, K, Kernytsky, A, Garimella, K, Altshuler, D, Gabriel, S, Daly, M, DePristo, MA (2010) The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, **20**(9): 1297–1303. eprint: 1011.1669.

Menardo, F, Loiseau, C, Brites, D, Coscolla, M, Gygli, SM, Rutaihwa, LK, Trauner, A, Beisel, C, Borrell, S, Gagneux, S (2018) Treemmer: a tool to reduce large phylogenetic datasets with minimal loss of diversity. *BMC Bioinformatics*, **19**(1): 164.

Miotto, P, Tessema, B, Tagliani, E, Chindelevitch, L, Starks, AM, Emerson, C, Hanna, D, Kim, PS, Liwski, R, Zignol, M, Gilpin, C, Niemann, S, Denkinger, CM, Fleming, J, Warren, RM, Crook, D, Posey, J, Gagneux, S, Hoffner, S, Rodrigues, C, Comas, I, Engelthaler, DM, Murray, M, Alland, D, Rigouts, L, Lange, C, Dheda, K, Hasan, R, Ranganathan, UDK, McNerney, R, Ezewudo, M, Cirillo, DM, Schito, M, Köser, CU, Rodwell, TC (2017) A standardised method for interpreting the association between mutations and phenotypic drug resistance in <>Mycobacterium tuberculosis<>. *European Respiratory Journal*, **50**(6): 1701354.

Paradis, E, Claude, J, Strimmer, K (2004) APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*, **20**(2): 289–290.

Pym, AS, Saint-Joanis, B, Cole, ST (2002) Effect of katG Mutations on the Virulence of Mycobacterium tuberculosis and the Implication for Transmission in Humans. *Infection and Immunity*, **70**(9): 4955–4960.

R Core Team (2013) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ramirez-Busby, SM, Valafar, F (2015) Systematic review of mutations in pyrazinamidase associated with pyrazinamide resistance in mycobacterium tuberculosis clinical isolates. *Antimicrobial Agents and Chemotherapy*, **59**(9): 5267–5277.

Ritter, C, Lucke, K, Sirgel, FA, Warren, RW, Helden, PD van, Bottger, EC, Bloemberg, GV (2014) Evaluation of the AID TB Resistance Line Probe Assay for Rapid Detection of Genetic Alterations Associated with Drug Resistance in Mycobacterium tuberculosis Strains. *Journal of Clinical Microbiology*, **52**(3): 940–946.

Safi, H, Lingaraju, S, Amin, A, Kim, S, Jones, M, Holmes, M, McNeil, M, Peterson, SN, Chatterjee, D, Fleischmann, R, Alland, D (2013) Evolution of high-level ethambutol-resistant tuberculosis through interacting mutations in decaprenylphosphoryl-β-D-arabinose biosynthetic and utilization pathway genes. *Nature Genetics*, **45**(10): 1190–1197.

Sirgel, FA, Wiid, IJ, Helden, PD van (2009) "Measuring Minimum Inhibitory Concentrations in Mycobacteria". In: *Mycobacteria Protocols. Methods in Molecular Biology (Methods and Protocols)*. Ed. by BA Parish T. Vol. 465. Totowa, NJ: Humana Press, pp. 173–186.

Springer, B, Lucke, K, Calligaris-Maibach, R, Ritter, C, Bottger, EC (2009) Quantitative Drug Susceptibility Testing of Mycobacterium tuberculosis by Use of MGIT 960 and EpiCenter Instrumentation. *Journal of Clinical Microbiology*, **47**(6): 1773–1780.

Stamatakis, A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **30**(9): 1312–1313. eprint: `bioinformatics/btu033` (10.1093).

Stucki, D (2015) Transmission and evolution of Mycobacterium tuberculosis studied by whole genome sequencing and single nucleotide polymorphism-typing.

Stucki, D, Brites, D, Jeljeli, L, Coscolla, M, Liu, Q, Trauner, A, Fenner, L, Rutaihwa, L, Borrell, S, Luo, T, Gao, Q, Kato-Maeda, M, Ballif, M, Egger, M, Macedo, R, Mardassi, H, Moreno, M, Tudo Vilanova, G, Fyfe, J, Globan, M, Thomas, J, Jamieson, F, Guthrie, JL, Asante-Poku, A, Yeboah-Manu, D, Wampande, E, Ssengooba, W, Joloba, M, Henry Boom, W, Basu, I, Bower, J, Saraiva, M, Vaconcellos, SEG, Suffys, P, Koch, A, Wilkinson, R, Gail-Bekker, L, Malla, B, Ley, SD, Beck, HP, Jong, BC de, Toit, K, Sanchez-Padilla, E, Bonnet, M, Gil-Brusola, A, Frank, M, Penlap Beng, VN, Eisenach,

K, Alani, I, Wangui Ndung'u, P, Revathi, G, Gehre, F, Akter, S, Ntoumi, F, Stewart-Isherwood, L, Ntinginya, NE, Rachow, A, Hoelscher, M, Cirillo, DM, Skenders, G, Hoffner, S, Bakonyte, D, Stakenas, P, Diel, R, Crudu, V, Moldovan, O, Al-Hajoj, S, Otero, L, Barletta, F, Jane Carter, E, Diero, L, Supply, P, Comas, I, Niemann, S, Gagneux, S (2016) Mycobacterium tuberculosis lineage 4 comprises globally distributed and geographically restricted sublineages. *Nature genetics*, **48**(12): 1535–1543.

Suarez, J, Ranguelova, K, Schelvis, JP, Magliozzo, RS (2009) Antibiotic resistance in Mycobacterium tuberculosis. Peroxidase intermediate bypass causes poor isoniazid activation by the S315G mutant of M. tuberculosis catalase-peroxidase (KatG). *Journal of Biological Chemistry*, **284**(24): 16146–16155.

Tudo, G, Rey, E, Borrell, S, Alcaide, F, Codina, G, Coll, P, Martin-Casabona, N, Montemayor, M, Moure, R, Orcau, A, Salvado, M, Vicente, E, Gonzalez-Martin, J (2010) Characterization of mutations in streptomycin-resistant Mycobacterium tuberculosis clinical isolates in the area of Barcelona. *Journal of Antimicrobial Chemotherapy*, **65**(11): 2341–2346.

Walker, TM, Ip, CL, Harrell, RH, Evans, JT, Kapatai, G, Dedicoat, MJ, Eyre, DW, Wilson, DJ, Hawkey, PM, Crook, DW, Parkhill, J, Harris, D, Walker, AS, Bowden, R, Monk, P, Smith, EG, Peto, TE (2013) Whole-genome sequencing to delineate Mycobacterium tuberculosis outbreaks: a retrospective observational study. *The Lancet Infectious Diseases*, **13**(2): 137–146.

Walker, TM, Kohl, TA, Omar, SV, Hedge, J, Del Ojo Elias, C, Bradley, P, Iqbal, Z, Feuerriegel, S, Niehaus, KE, Wilson, DJ, Clifton, DA, Kapatai, G, Ip, CLC, Bowden, R, Drobniewski, FA, Allix-Béguec, C, Gaudin, C, Parkhill, J, Diel, R, Supply, P, Crook, DW, Smith, EG, Walker, AS, Ismail, N, Niemann, S, Peto, TEA (2015) Whole-genome sequencing for prediction of Mycobacterium tuberculosis drug susceptibility and resistance: a retrospective cohort study. *The Lancet Infectious Diseases*, **15**(10): 1193–1202.

Yu, G, Smith, DK, Zhu, H, Guan, Y, Lam, TTY (2017) ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*, **8**(1): ed. by G McInerny, 28–36.

Zaunbrecher, MA, Sikes, RD, Metchock, B, Shinnick, TM, Posey, JE (2009) Overexpression of the chromosomally encoded aminoglycoside acetyltransferase eis confers kanamycin resistance in Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences*, **106**(47): 20004–20009.

B. List of publications

Gygli, S. M., Borrell, S., Trauner, A., & Gagneux, S. (2017). Antimicrobial resistance in Mycobacterium tuberculosis: mechanistic and evolutionary perspectives. FEMS Microbiology Reviews, 41(3), 354–373. doi:10.1093/femsre/fux011

Gygli, S. M., Keller, P. M., Ballif, M., Blöchliger, N., Hömke, R., Reinhard, M., Böttger, E. C. (2018). Whole genome sequencing for drug resistance profile prediction in Mycobacterium tuberculosis. BioRxiv, 401703. doi:https://doi.org/10.1101/401703

Menardo, F., Loiseau, C., Brites, D., Coscolla, M., **Gygli, S. M**., Rutaihwa, L. K., Gagneux, S. (2018). Treemmer: a tool to reduce large phylogenetic datasets with minimal loss of diversity. BMC Bioinformatics, 19(1), 164. doi:10.1186/s12859-018-2164-8

Rutaihwa, L. K., Menardo, F., Stucki, D., **Gygli, S. M.**, Ley, S. D., Malla, B., Gagneux, S. (2018). Multiple introductions of the Mycobacterium tuberculosis Lineage 2 Beijing into Africa over centuries. BioRxiv, 413039. doi:doi.org/10.1101/456434

Trauner, A., Banaei-esfahani, A., **Gygli, S.**, & Warmer, P. (2018). Resource misallocation as a mediator of fitness costs in antibiotic resistance. BioRxiv, 456434. doi:10.1101/456434

Wittwer, M., Altpeter, E., Pilo, P., **Gygli, S. M.**, Beuret, C., Foucault, F., Schürch, N. (2018). Population Genomics of Francisella tularensis subsp. holarctica and its Implication on the Eco-Epidemiology of Tularemia in Switzerland. Frontiers in Cellular and Infection Microbiology, 8(March), 1–16. doi:10.3389/fcimb.2018.00089